Targeting the pupylation cascade system of *Mycobacterium tuberculosis* for the development of novel anti-tubercular drugs

by

SANDILE M. MTHEMBU

Submitted in full fulfilment of the academic requirements of

Doctor of Philosophy

in Biochemistry

School of Life Sciences

College of Agriculture, Engineering and Science

University of KwaZulu-Natal

Pietermaritzburg

South Africa

12 December 2025

PREFACE

The research contained in this thesis was completed by the candidate while based in the Discipline of Biochemistry, School of Life Sciences of the College of Agriculture, Engineering and Science, University of KwaZulu-Natal, South Africa. The research was financially supported by National Research Foundation (NRF; UID: 140583).

The contents of this work have not been submitted in any form to another university and, except where the work of others is acknowledged in the text, the results reported are due to investigations by the candidate.



Signed: Supervisor

Date: 18 December 2025



Signed: Co-supervisor

Date: 18 December 2025

DECLARATION 1: PLAGIARISM

Note that two declaration sections are required if there are papers emanating from the dissertation/thesis. The first (obligatory) declaration concerns plagiarism and the second declaration specifies your role in the published papers.

I, Sandile Mehzi Mthembu, declare that:

(i) the research reported in this dissertation, except where otherwise indicated or acknowledged, is my original work;

(ii) this dissertation has not been submitted in full or in part for any degree or examination to any other university;

(iii) this dissertation does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons;

(iv) this dissertation does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:

a. their words have been re-written but the general information attributed to them has been referenced;

b. where their exact words have been used, their writing has been placed inside quotation marks, and referenced;

(v) where I have used material for which publications followed, I have indicated in detail my role in the work;

(vi) this dissertation is primarily a collection of material, prepared by myself, published as journal articles or presented as a poster and oral presentations at conferences. In some cases, additional material has been included;

(vii) this dissertation does not contain text, graphics or tables copied from the Internet, unless specifically acknowledged, and the source being detailed in the dissertation and in the References sections.

(viii) this dissertation does not contain text, graphics or tables generated and/or copied using Artificial Intelligent systems (such as ChatGPT, Google Gemini. Cortana, Writesonic, Copy.ai etc).



Signed: Sandile Mehzi Mthembu

Date: 12 / 12 / 2025

DECLARATION 2: PUBLICATIONS

Chetty, R., Delport, A., Mthembu, S., Veale, C.G*. and Hewer, R*., 2025. Screening the Pandemic Response Box identifies novel ligands of the Staphylococcus aureus protein arginine kinase, McsB. *Molecular Biology Reports*, *52*(1), p.446.

Author contributions (**CRediT**)

The study was conceived and designed by Clinton Veale and Raymond Hewer, who also provided supervision and secured funding. Material preparation, data collection, and formal analysis were carried out by Ryan Chetty, Alexandré Delport, and Sandile Mthembu. The original draft of the manuscript was written by Ryan Chetty and Alexandré Delport, with subsequent drafts reviewed and edited by Clinton Veale and Raymond Hewer.

Presentations

Mthembu S, Delport A, Hewer, R*. *Targeting the pupylation cascade system of Mycobacterium tuberculosis for the development of novel anti-tubercular drugs. South African Society of Biochemistry and Molecular Biology*, 7 July 2024, Polokwane, South Africa. Oral presentation by S. Mthembu

Mthembu S, Delport A, Hewer, R*. *Targeting the pupylation cascade system of Mycobacterium tuberculosis for the development of novel anti-tubercular drugs. UKZN's Postgraduate Research and Innovation Symposium,* 29 October 2024, Durban, South Africa. Oral presentation by S. Mthembu

Mthembu S, Delport A, Hewer, R*. *Targeting the pupylation cascade system of Mycobacterium tuberculosis for the development of novel anti-tubercular drugs. UKZN's Postgraduate Seminar series,* 20 November 2024, Durban, South Africa. Oral presentation by S. Mthembu



Signed: Sandile Mehzi Mthembu

Date: 12 / 12 / 2025

ABSTRACT

Tuberculosis (TB) remains the world's deadliest infectious disease, driven by the ability of *Mycobacterium tuberculosis* (*Mtb*) to establish chronic infection within alveolar macrophages. The global burden is exacerbated by the rise of antibiotic-resistant strains, underscoring an urgent need for novel therapeutic strategies and drug targets. The mycobacterial Pup-proteasome system (PPS), a bacterial degradation pathway essential for stress response and virulence, represents a promising yet underexplored target for anti-mycobacterial drug development.

This study focused on the PPS ligase, proteasome accessory factor A (PafA), as a key therapeutic target. Recombinant *Mtb* PafA was produced, purified to >95% homogeneity, and biophysically characterised. Notably, *Mtb* PafA exists as a constitutively active monomer—distinct from the described dimeric forms of other PafA homologs—with Mg^{2+} acting as an allosteric regulator of the PPS system. Specifically, Mg^{2+} modulates PafA activity by inducing concentration-dependent conformational changes that govern ATP-PafA interactions, establishing a previously unrecognised allosteric regulatory mechanism within the PPS. Optimised thermal shift assay (TSA) and cellular thermal shift assay (CETSA) platforms were developed to screen and validate a subset of the ChemBridge DiverSET chemical library. From this screen, seven molecules were validated to bind soluble *Mtb* PafA, including C10 (ΔT_m = 1.10°C; ΔTagg = 10.29°C; docking = -8.474 kcal/mol), A2 (ΔT_m = 1.20°C; ΔTagg = 8.60°C; docking = -6.440 kcal/mol), D2 (ΔT_m = 0.75°C; ΔTagg = 8.48°C; docking = -8.592 kcal/mol), and A3 (ΔT_m = 0.85°C; ΔTagg = 9.31°C; docking = -8.338 kcal/mol)—values comparable to those of the natural substrate ATP (ΔT_m = 1.33–2.22°C; docking = -11.278 kcal/mol). Although computational predictions suggested competitive binding to the ATP-binding pocket of PafA, subsequent growth assays revealed no significant effect on mycobacterial proliferation under oxidative stress conditions, indicating a non-inhibitory, alternative ligand role within the PPS rather than direct antibacterial activity.

These findings establish PafA as a ligandable target, addressing a critical gap in anti-mycobacterial therapeutics by expanding the druggable landscape to include proteostasis pathways. This offers a promising translational trajectory toward new agents active against drug-resistant tuberculosis by establishing novel biochemically validated PafA-binding scaffolds that require further optimization.

ACKNOWLEDGMENTS

I extend my deepest gratitude to my supervisors, Prof. R. Hewer and Dr. A. Delport. Their expert guidance, profound knowledge, and insightful advice provided the intellectual foundation for this work, giving me the confidence to navigate and overcome all the challenges encountered throughout the study. I am sincerely thankful for their unwavering commitment and patience, which were instrumental in bringing this research to fruition. Your dedication has profoundly shaped not only my scientific career but also those of the entire team, fostering a culture of excellence.

My sincere thanks also go to my colleagues in Lab 12 (formally 46) for their steadfast support, camaraderie, and encouragement throughout this journey, which made each and every workday both productive and enjoyable. I also wish to thank my peers in the Block B office for their motivation, constructive discussions, and shared perseverance.

I acknowledge, with gratitude, the financial support provided by the National Research Foundation and the University of KwaZulu-Natal for the laboratory facilities and resources essential to this research, as well as all the experiences that arose through the institution's various programs, which significantly contributed to my overall growth.

Lastly, and most importantly, I would like to express my eternal gratitude to my family for their unwavering love and support. To my mother, Mrs. B.R. Gwija, my grandmama, Mrs. B.R. Mthembu, and my older and younger sisters—your belief in me has been my anchor. I also pay tribute to the memory of my dad, Mr. V. Gwija, whose presence I deeply missed during the course of this study.

TABLE OF CONTENTS

PREFACE ii
DECLARATION 1: PLAGIARISM iii
DECLARATION 2: PUBLICATIONS iv
ABSTRACT vii
ACKNOWLEDGMENTS viii
TABLE OF CONTENTS ix
LIST OF FIGURES xi
LIST OF TABLES xiv
LIST OF ABBREVIATIONS xv
CHAPTER 1. LITERATURE REVIEW 1
1.1. Introduction 1
1.2 Mycobacterial treatment and emergence of drug resistance 2
1.3 Mycobacterial evolution and threat to global health security 5
1.4 Novel drug development strategies to combat existing and emerging drug-resistant mycobacterial strains 8
1.4.1 The mycobacterial PPS system 9
1.4.2 PPS role in stress response 11
1.4.3 PPS role in aiding mycobacterial pathogenicity 12
1.4.4 PPS as an anti-mycobacterial target 19
1.5 Study rationale 23
1.6 Novelty 23
1.7 Significance and impact 24
CHAPTER 2. MATERIAL AND METHODS 25
2.1. Recombinant expression and purification of pupylation proteins 25
2.1.1 Competent cell preparation and transformation 25
2.1.2 GroEL-GroES assisted recombinant co-expression of PafA 26
2.1.3 Recombinant PafA purification 27
2.1.4 Polyacrylamide gel electrophoresis and western blot analysis 28
2.2 PafA ligandability studies 29
2.2.1 PafA TSA optimization 29
2.2.2 PafA TSA: Small-molecule ligand screening 30
2.2.3 PafA CETSA: *in vivo* PafA ligandability studies 31

2.3 In-Silico characterization of PafA binders....32
2.3.1 Sequencing analysis of *Mtb* PPS proteins against resolved PPS homolog models....32
2.3.2 In vitro PafA ligands docking analysis....33
2.3. Effect of *In vitro* ligands in the mycobacterial PPS....34
2.3.1. *In vitro* ligands effect under standard growth conditions....34
2.3.2. *In vitro* ligands effect under oxidative stress growth conditions....34
CHAPTER 3. RESULTS....36
3.1 Recombinant expression and purification of pupylation proteins....36
3.1.1 GroEL-GroES-assisted expression of recombinant PafA....36
3.1.2 Recombinant PafA purification....37
3.1.3 PafA biophysical characterization....39
3.2 Identification and biophysical validation of novel PafA ligands....45
3.2.1 Identifying novel PafA ligands via the thermal shift assay....45
3.2.2 *In vitro* ligands *in silico* validation....58
3.2.3 PafA ligands validation via the cellular thermal shift assay (CETSA)....78
3.3. Effect of *In vitro* ligands in the mycobacterial PPS....85
3.3.1. *In vitro* ligands effect on mycobacterial proliferation....85
3.3.2. *In vitro* ligand effects under oxidative stress growth conditions....87
CHAPTER 4. DISCUSSION....89
4.1 *Mtb* PafA is a constitutively active monomer, streamlining pupylation activation....89
4.2 Mg^{2+} Acts as an Allosteric Activator of the Monomeric *Mtb* PafA for enzyme-level regulation 92
4.3 PafA is a ligandable target with high structural plasticity....94
REFERENCES....100
APPENDIX....111

LIST OF FIGURES

Figure 1.1 A proposed evolutionary model for host-adapted mycobacteria. 6

Figure 1.2. A proposed model for the adaptation of environmental mycobacteria to intracellular life. 7

Figure 1.3. Schematic representation of the Pup-proteasomal degradation cascade system (PPS) in mycobacteria..... 9

Figure 1.4. Dop and PafA structural homology. 10

Figure 1.5 The PafABC operon in *Mycobacterium tuberculosis* (*Mtb*). 13

Figure 1.6 PafABC operon role in the pup proteasome degradation cascade system (PPS). 14

Figure 1.7. PafBC-mediated reprogramming of RNA polymerase in mycobacterial DNA damage response.....15

Figure 1.8. Coordination and regulation of DNA-damage repair by the PafBC and SOS pathways. 16

Figure 1.9 The hypothetical DopPupPcrAB operon in *Mtb*. 17

Figure 1.10 Schematic of PROTAC-mediated targeted protein degradation......21

Figure 3.1 Soluble co-expression of recombinant PafA with the GroEL in *E. coli.* 36

Figure 3.2. **Small-scale recombinant PafA purification A**. 38

Figure 3.3. Scaled-up recombinant PafA purification.....39

Figure 3.4. Recombinant PafA molecular exclusion chromatography (SEC) elution profile. 40

Figure 3.5. Analysis of recombinant PafA fractions from SEC. 41

Figure 3.6. Native-state detection of recombinant *Mtb* PafA. 42

Figure 3.7. PafA Sephadex G75 SEC purification.. 43

Figure 3.8. **Sephadex G75 SEC purification**. 44

Figure 3.9. **Thermal melt analysis of purified PafA.**. 45

Figure 3.10. Thermal melt analysis of stabilised PafA. 46

Figure 3.11. PafA stability buffer optimization.. 47

Figure 3.12. PafA oligomeric state examination via the thermal melt assay. 48

Figure 3.13. Effect of Mg^{2+} and ATP on PafA thermal stability. 49

Figure 3.14. Thermal stabilization of PafA by ATP and Mg^{2+} 51

Figure 3.15 PafA thermal melt profile in the presence of bithionol.....52

Figure 3.16. PafA thermal melt profile in the presence of ST1926. 53

Figure 3.17. The identification of novel PafA ligands via the PafA TSA 56

Figure 3.18. PafA thermal stability in the presence of small-molecule ligands. 57

Figure 3.19. Structural conservation analysis of actinobacterial PafA homologs.....59

Figure 3.20. Architectural comparison of the cgPafA ATP and Pup-binding sites.....62

Figure 3.21. Comparative stability of PafA constructs with and without native ligands 65

Figure 3.22. Comparative stability of cgPafA construct bound by native and *in vitro* ligands. ...66

Figure 3.23. Compound A3 docking cgPafA ATP site......68

Figure 3.24 Compound C10 docking cgPafA ATP site. 69

Figure 3.25 Compound D2 docking cgPafA ATP site. 70

Figure 3.26 Comparative stability of cgPafA construct in complex with ligand A3 with and without Mg^{2+} 72

Figure 3.27 Compound A3 docking cgPafA ATP site in the absence of Mg^{2+} 73

Figure 3.28. Compound A3 interactions within cgPafA ATP site in the absence of Mg2+ 74

Figure 3.29 Compound C9 docking the mtPafA Pup^{IQK} groove site. 75

Figure 3.30 Comparative stability of cgPafA construct in complex with ligand C9 77

Figure 3.31 Compound C9 docking the mtPafA Pup^{IQK} groove site......77

Figure 3.32 Ponceau S Stain of PafA CETSA samples......78

Figure 3.33. Recombinant PafA cellular thermal shift assay (CETSA) optimisation. Recombinant PafA lysates were incubated at a gradient of temperatures.....79

Figure 3.34. Control PafA CETSA. 80

Figure 3.35. Recombinant PafA ΔT_{agg} in the presence of control ligands. 81

Figure 3.36. PafA cellular thermal shift assay (CETSA) in the presence of the identified PafA binders.....83

Figure 3.37. Thermal stabilization of PafA by small-molecule drug-like compounds. 84

Figure 3.38. PafA *in vitro* ligands effect on mycobacterial cell proliferation.....85

Figure 3.39. Nitrosative stress mycobacterial proliferation assay optimisation.. 87

Figure 3.40. PafA in vitro ligands effect on mycobacterial cell proliferation under oxidative stress conditions.....88

LIST OF TABLES

Table 1.1. Summary of *Mtb* mutated genes conferring resistance against first-line defence anti-*Mtb* drugs 4

Table 2.1. Preparation of PafA TSA reactions 30

Table 3.1. Optimization of the PafA thermal shift assay using native ligands 50

Table 3.2. PafA thermal stability analysis in the presence of controls bithionol and ST1926 54

Table 3.3. PafA thermal stability in the presence of selected small-molecule ligands.....58

Table 3.4. Ligand docking scores and MM/GBSA free-binding energies for *in vitro* PafA ligands docked into the ATP site with and without Mg^{2+} of PafA from *Corynebacterium glutamicum* 60

Table 3.5. Ligand docking scores and MM/GBSA free-binding energies for *in vitro* PafA ligands docked into the Pup binding groove of PafA from *Corynebacterium glutamicum* 63

Table 3.6. Ligand docking scores and MM/GBSA free-binding energies of ST1926 and *in vitro* mtPafA ligands docked into the Pup-binding groove of PafA. 64

Table 3.7. RMSD, Radius of Gyration, RMSF and RMSD of ligand only of cgPafA and ligands docked into the ATP site after 10-100 ns MD simulations.....67

Table 3.8. Mean free binding energies of cgPafA complexes post molecular dynamics simulations 71

Table 3.9. RMSD, Radius of Gyration, RMSF and RMSD of ligand only of mtPafA and mtPafA-C9 complex after 10-100 ns MD simulations 76

Table 3.10. PafA CETSA analysis in the presence of the identified PafA ligands 82

Table 3.11. PafA CETSA analysis in the presence of the identified PafA ligands 84

Table 3.12. PafA *in vitro* ligands effect on *Msm* cell proliferation under standard conditions. ...86

Table 4.1. Cross assay binding activity of PafA ligands. 99

LIST OF ABBREVIATIONS

Abbreviation	Definition
1xEB	Equilibration buffer
Ace	*Acidothermus cellulolyticus*
acDop	*Acidothermus cellulolyticus Dop*
AEBSF	4-(2-Aminoethyl)benzenesulfonyl fluoride hydrochloride
ATP	Adenosine triphosphate
BacPROTAC	Bacterial PROTAC
BRD1	Bromodomain 1
BSA	Bovine serum albumin
CETSA	Cellular thermal shift assay
Cgl	*Corynebacterium glutamicum*
cgPafA	*Corynebacterium glutamicum* PafA
Clp	Caseinolytic protease
CymA	Cyclomarin A
DDR	DNA-damage response
DMSO	Dimethyl sulfoxide
Dop	Deamidase of Pup
DTT	Dithiothreitol
ECL	Enhanced chemiluminescence
ELISA	Enzyme-Linked Immunosorbent Assay
FabD	Malonyl-CoA-acyl carrier protein transacylase
HGT	Horizontal gene transfer
HRP	Horse radish peroxidase
IFN- γ	Interferon-γ
IMAC	Immobilized metal affinity chromatography
LB	Luria-Bertani medium
Maf	*Mycobacterium africanum*
Mas	*Mycobacterium abscessus*
Mav	*Mycobacterium avium*
Mbv	*Mycobacterium bovis*
Mcp	*Mycobacterium caprae*
Mct	*Mycobacterium canettii*
MD	Molecular dynamics
Mic	*Mycobacterium intracellulare*

Mks	*Mycobacterium kansasii*
Mmc	*Mycobacterium microti*
Mmg	*Mycobacterium Mungi*
MM/GBSA	Molecular Mechanics Generalized Born Surface Area
Mog	*Mycobacterium orygis*
Mpa	Mycobacterial proteasome ATPase
Mpp	*Mycobacterium Pinnipedii*
mSA	Monomeric streptavidin
Msm	*Mycobacterium smegmatis*
Mtb	*Mycobacterium tuberculosis*
MTBC	*Mycobacterium tuberculosis* complex
mtDop	Mutated *Mtb* Dop
mtPafA	Mutated *Mtb* PafA
NTM	Non-tubercular mycobacteria
PafA	Proteasome accessory factor A
PafB	Proteasome accessory factor B
PafC	Proteasome accessory factor C
PanB	3-methyl-2-oxobutanoate hydroxymethyltransferase
PPS	Pup-proteasome protein recycling cascade system
prcA	Proteasome subunit alpha
prcB	Proteasome subunit beta
PROTAC	Proteolysis targeting chimera
$Pup^{Q/E}$	Prokaryotic ubiquitin-like protein $^{\text{glutamine/glutamate c-terminal residue}}$
RecA	Recombinase A
RMSD	Root mean square deviation
RMSF	Root mean square fluctuation
RoGyr	Radius of gyration
RT	Room temperature
SDS	Sodium dodecyl sulfate
SEC	Size exclusion chromatography
SR	Stringent response
ssDNA	Single-stranded DNA
TB	Tuberculosis
TBS	Tris-buffered saline
TSA	Thermal shift assay
Tm	Protein melting temperature

UV	Ultraviolet
WHO	World Health Organization
YOPI	Young, old, pregnant, and immunosuppressed

CHAPTER 1. LITERATURE REVIEW

1.1. Introduction

TB, a disease with a nine-millennia-long association with humanity, remains a major concern despite significant advances in its epidemiology (Kestler and Tyler, 2022). Human pulmonary TB is primarily caused by prolonged infection with *Mtb*; however, other human-adapted species within the zoonotic *Mycobacterium tuberculosis* complex (*MTBC*) can also be responsible (Kock *et al*., 2021). The *MTBC* comprises a group of nine closely-related species of mycobacteria that includes *Mtb*, *Mycobacterium bovis* (*Mbv*), *Mycobacterium africanum* (*Maf*), *Mycobacterium microti* (*Mmc*), *Mycobacterium caprae (Mcp)*, *Mycobacterium orygis (Mog)*, *Mycobacterium pinnipedii (Mpp)*, *Mycobacterium mungi (Mmg)*, and *Mycobacterium canettii* (*Mct*) (Brites and Gagneux, 2017). Although *MTBC* members demonstrate exceptional genomic similarity (>99%), phenotypic variation is sufficient to drive significantly divergent immunological responses, culminating in pulmonary TB disease (Zwyer *et al.*, 2023). The specific adaptations of each *MTBC* member to their respective host accounts for this variability. The spectrum of host tropism is wide: *Mtb* and the West African endemic *Maf* are human-adapted, *Mbv* and *Mcp* target cattle and ruminants, *Mmc* is specialized for voles, pinnipeds as hosts for *Mpp*, *Mmg* infecting banded mongooses, *Mog* has a broad range of hosts including cattle and wildlife, while *Mct* has no identified primary host (Niemann and Supply, 2014).

TB consistently ranks as the world's deadliest infectious disease, a position it reclaimed following the SARS-CoV-2 pandemic (Gulumbe *et al*., 2025). Current estimates indicate that approximately 25% of the global population shows an immunological response to TB infection, which can remain a latent, dormant state or progress into active disease (Goletti *et al*., 2025). *Mtb*-mediated TB infection is typically transmitted when a person inhales aerosolized droplets from an individual with active TB disease. Upon inhalation, *Mtb* travels to the alveolar sacs, multiplies, and triggers an immune response that culminates in its phagocytosis by IFN-γ-activated macrophages (Queval *et al*., 2017). Although this process typically mounts an effective host defence by promoting the destruction of foreign bacteria, it is paradoxically a key evolutionary component of *MTBC*'s obligatory pathogenicity. Through the inhibition of phagosome-lysosome fusion, *Mtb* halts macrophage maturation thereby evading the innate immune response and effectively commandeers the host cell for its replication (Queval *et al*.,

2017). This intracellular persistence triggers the recruitment of other immune cells to the infection site, which differentiate and aggregate to form a granuloma—the pathological hallmark of TB.

The fate of *Mtb* cells depends on host immunity. While the bacilli often lie dormant within granuloma macrophages for years in immunocompetent hosts without causing disease, resulting in latent infection, they can reactivate in vulnerable YOPI populations (young, old, pregnant, and immunosuppressed). In these individuals, the bacteria replicate, breach the granuloma, and disseminate to other lung regions, propagating active TB disease (Ahmad, 2011). The escape of replicated *Mtb* from the primary infection site also allows the infection to spread to distant tissues and organs—such as the spine, kidneys, or brain—via the lymphatic and hematologic systems, resulting in the complicated condition of disseminated TB. The most severe form of this is miliary TB, an extremely rare disease defined by the widespread, simultaneous infection of multiple organ systems (Mohammadnabi *et al*., 2024).

Despite being a preventable and curable disease, TB continues to cause millions of deaths and new infections annually, constituting a major public health paradox. The root of this contradiction lies in socio-economic determinants of health, primarily poverty, that obstruct clinical care and amplify transmission risks. Extensive research has demonstrated that poverty indirectly contributes to TB prevalence by systematically restricting access to preventive education, timely diagnosis, and effective treatment (Ahn and WHO, 2005; Jannsens and Rieder, 2008; Oxlade and Murray, 2012; Reis-Santos *et al.*, 2019; Murdoch *et al.*, 2021). Compounding this, non-adherence to the necessary lengthy drug therapies promotes the emergence and dissemination of drug-resistant mycobacterial strains, a current topic of major discussion.

1.2 Mycobacterial treatment and emergence of drug resistance

Typically associated with inhibitory drugs, drug resistance has been a surging epidemic that limits pre-existing clinical interventions - threatening global health security. As one of the most significant discoveries to have shaped modern medicine, antibiotics play a major role in clinical intervention of a range of bacterial infections, including TB disease. Initial TB treatment against *Mtb* often employs a standard FDA-approved 4-drug regimen inclusive of rifampin, isoniazid, pyrazinamide, and ethambutol; the latter being immediately suspended upon confirmation of

treatment efficacy. Drug-susceptible *Mtb* is typically treated with a six-month regimen: a two-month intensive phase using all four first-line drugs, followed by a further four-month continuation phase of rifampicin and isoniazid (Nahid *et al*., 2016). Shorter regimens are currently under review to help combat the rise of resistant *Mtb* strains, with a recent study by the Centers for Disease Control and Prevention in collaboration with the AIDS Clinical Trials Group outlining a promising new 4-month treatment regimen using rifapentine, isoniazid, pyrazinamide, and moxifloxacin (Dorman *et al*., 2021). According to the TB Trials Consortium study 31/A5349, a 4-month regimen of rifapentine and moxifloxacin, which replaced rifampin and ethambutol, demonstrated non-inferior efficacy compared to the standard 6-month regimen for treating drug-susceptible *Mtb*. Importantly, this comparable efficacy was consistent in patients both with and without HIV co-infection (Dorman *et al*., 2021; Pettit *et al.*, 2023).

Unfortunately, the global antimicrobial resistance epidemic is increasingly limiting therapeutic options for TB, posing a severe challenge to its clinical management. The situation is evidenced by the emergence of *Mtb* strains with varying resistance profiles, which range from resistance to a single first-line agent to extensively and even totally drug-resistant strains that are impervious to all standard drug therapies (Shah *et al.*, 2005; Velayati *et al.*, 2009). The challenge of drug resistance extends beyond tuberculosis, compromising antibiotic treatment for a range of illnesses, with leprosy serving as a critical example. Caused by *Mycobacterium leprae* or *Mycobacterium lepromatosis,* leprosy is a chronic infection that destroys peripheral nerves and skin, which—if left untreated—causes significant permanent deformities and disability. Like TB, leprosy treatment using the World Health Organization (WHO) recommended multidrug treatment (Rifampicin, Dapsone, and Oxfloxacin) is also jeopardized by antimicrobial resistance, a phenomenon commonly achieved through mutation-based mechanisms that modulate therapeutic target sites (*rpoβ*, *FolP1*, and *GyrA*) (Sekar *et al*., 2011; Perera *et al*., 2025).

Evidently, the primary mechanism of drug resistance in mycobacteria involves mutations that alter the drug target sites, often reducing or eliminating a drug's binding affinity, thereby nullifying its inhibitory potential. The seminal work of Telenti and collaborators was the first to characterize this mechanism in the context of TB antimicrobial resistance (Telenti *et al*., 1993). Through sequencing analysis, the group revealed that high-level rifampicin resistance in *Mtb* was caused by mutations in the *rpoβ* gene. This gene encodes the β-subunit of the DNA-dependent RNA polymerase, and the resistance-conferring mutations specifically cluster in

eight conserved amino acids (Telenti *et al*., 1993). As new resistant bacterial strains continually emerge, a growing number of mutation-based resistance mechanisms against various clinical drug targets are documented; Table 1.1 summarizes several classical examples.

Table 1.1. Summary of *Mtb* mutated genes conferring resistance against first-line defence anti-*Mtb* drugs

Drug	Mutated gene for high-level resistance
Rifampin	*rpoB*
Isoniazid	*katC*, *inhA*, *oxyR*, *ahpC*, *furA*
Pyrazinamide	*pncA*, IS*6110* insertion
Ethambutol	*embB*
Streptomycin	*rrs, rpsL*
Dapsone	*FolP1*
Oxfloxacin	*GyrA*

(Gillespie, 2002; Lavania *et al*., 2018; Li *et al*., 2012; Li *et al*., 2022*)

The drug resistance epidemic is also fuelled by horizontal gene transfer (HGT), a process that allows most bacteria to share genetic material, including resistance genes, within and between species. Although members of the *MTBC* are considered extremely clonal and long thought to evolve primarily through clonal expansion, in contrast to HGT-influenced speciation, emerging evidence indicates that HGT has significantly contributed to the evolution of *MTBC'*s virulence factors (Veyrier *et al.*, 2011; Boritsch and Brosch, 2016; Gagneux, 2018; Bryant *et al.*, 2021). Comprehensive reviews by Veyrier *et al*. (2011) and Gagneux. (2018) have detailed the evolutionary origins of the *MTBC*, particularly highlighting *Mtb'*s potential biphasic evolution which involved both recombination via HGT and genome reduction events.

The most notable NTM pathogens—including *Mycobacterium avium* (*Mav*), *Mycobacterium intracellulare* (*Mic*), *Mycobacterium kansasii* (*Mks*), and the particularly concerning *Mycobacterium abscessus* (*Mas*)—typically cause chronic pulmonary disease resembling tuberculosis, in addition to other cutaneous and pulmonary illnesses. *Mas* is especially problematic. This rapidly growing, multidrug-resistant NTM frequently contaminates medical instruments, causing intractable infections that complicate clinical therapies. Despite the absence of conclusive evidence for HGT in the *MTBC* and the prevailing model of mutation-based evolution, the possibility of rampant HGT events between NTM and the *MTBC* has remained a significant concern, as it could account for the rapid acquisition of drug resistance

in clinical *Mtb.* A 2023 genetic study by Xia directly challenged this concern by providing evidence that while HGT events are indeed present in *Mtb*, such transfers are restricted to non-essential genes (Xia, 2023). These findings reinforce the conclusion that drug resistance in the *Mtb,* or rather in *MTBC* members, is primarily mediated by mutations.

Unlike in *MTBC*, HGT events are widespread in NTM and are a significant mediator of drug resistance. This is exemplified by the work of Nguyen and collaborators who observed the routine conjugal transfer of chromosomal and extrachromosomal DNA between *M. smegmatis* strains on both solid medium and within complex biofilm communities, significantly influencing mycobacterial adaptive evolution through genomic diversity (Nguyen *et al.*, 2010). Plasmid DNA transfer is a common HGT mechanism in fast-growing NTM that readily spreads drug resistance genes. Despite their potential clinical importance, very few studies have explored these plasmids and the functionality of the genes they encode. However, a recent analysis of the NTM 'plasmidome'—which included clinically relevant isolates—revealed that these plasmids are both widespread and highly diverse (Diricks *et al.*, 2025).

Notably, among key pathogens such as *Mav*, *Mic*, *Mks*, *Mas,* and *Mycobacterium ulcerans*, significant variation in the plasmids was observed not only between species but also in their genetic cargo. Critically, some plasmids were found to carry genes that confer resistance to drugs commonly used to treat NTM infections (Diricks *et al.*, 2025). The prevalence of plasmids in NTM stands in stark contrast to their exceptional rarity in the *MTBC*, where initial, tentative evidence by Zainuddin and Dale (1990) reporting the presence of low-level presumptive extrachromosomal DNA in *Mtb,* was only definitively corroborated decades later by Ummels *et al*. (2014). The latter study provided the first conclusive evidence of potential conjugative HGT in the complex by identifying a unique plasmid transmissible among slow-growing mycobacteria, including *Mtb* and *Mbv* (Ummels *et al.*, 2014).

1.3 Mycobacterial evolution and threat to global health security

Importantly, *MTBC* ecology and evolutionary exploration data suggest origins from an environmental NTM-like opportunistic pathogen ancestor before the split event that categorizes mycobacteria into the two major current groups, i.e., the slow- and fast-growing mycobacteria. Specifically, genomic evidence indicates the *MTBC* diverged from environmental, slow-growing mycobacterial ancestors via a biphasic evolution: the acquisition of virulence genes

through HGT recombination events, followed by genome reduction. These complementary events collectively promoted an intracellular lifestyle, resulting in the obligate pathogenicity that defines the complex, a trait optimized for host replication and transmission (Figure 1.1).

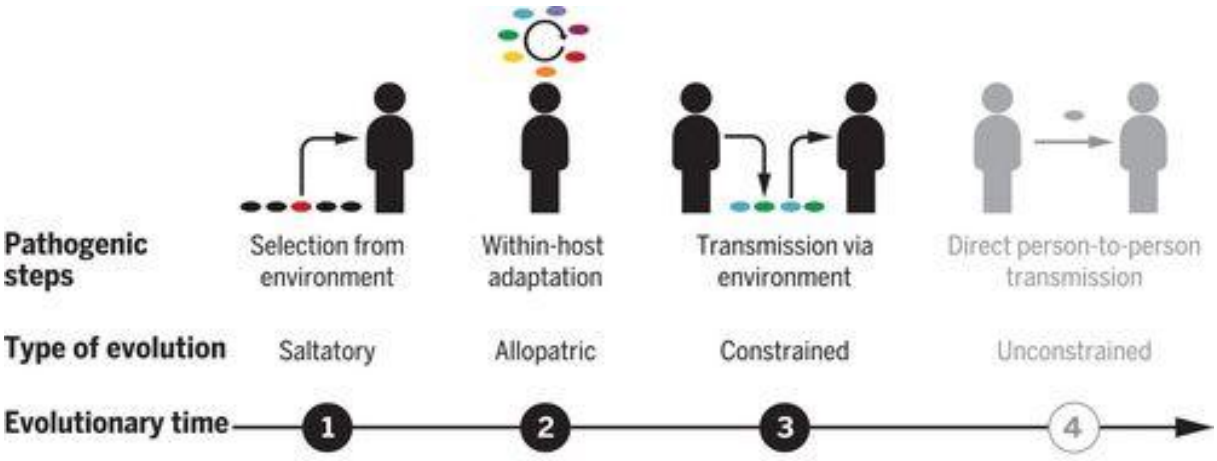

Figure 1.1 A proposed evolutionary model for host-adapted mycobacteria. The model outlines a potential evolutionary trajectory from environmental, non-tuberculous mycobacteria (NTM) to obligate pathogens of the *Mycobacterium tuberculosis* complex (*MTBC*). (1) Ancestral NTM undergo horizontal gene transfer (HGT), acquiring virulence factors that facilitate survival in environmental amoebae and, incidentally, in macrophages upon sporadic host encounter. (2) Chronic infection within a mammalian host drives allopatric evolution, refining intracellular adaptations. A critical bottleneck arises at transmission: (3) Most host-adapted lineages become constrained, losing the capacity for efficient host-to-host spread and therefore are unable to spread. (4) In rare instances, an adapted lineage overcomes this constraint, evolving efficient transmission mechanisms to become an obligate, host-restricted pathogen akin to modern MTBC members. (Bryant *et al*., 2021).

A widely accepted theory posits that the evolution of mycobacteria toward obligate pathogenicity occurred in two stages: the initial intracellular adaptation within environmental amoeba, followed by further specialization during mammalian host infection (Price *et al*., 2024). This model is strongly supported by the mechanistic parallels in how both eukaryotic cell types process bacteria—through phagocytosis and phagolysosome formation—coupled with the documented ecological overlap and direct interactions between amoeba and mycobacteria. Moreover, several mycobacterial species from both the *MTBC* and NTM groups, including *Mtb*, are phagocytosed by amoebae and persist intracellularly by deploying similar tactics used to survive within mammalian host macrophages, such as inhibiting phagosome-lysosome fusion (Figure 1.2). This remarkable parallel positions amoebae as a crucial environmental training ground, where mycobacteria evolve the ability to thrive inside a host cell by evading defences and commandeering resources—a skill set directly applied in mammalian infection. Consequently, this interaction is fundamental to the survival, transmission, and evolution of mycobacterial virulence.

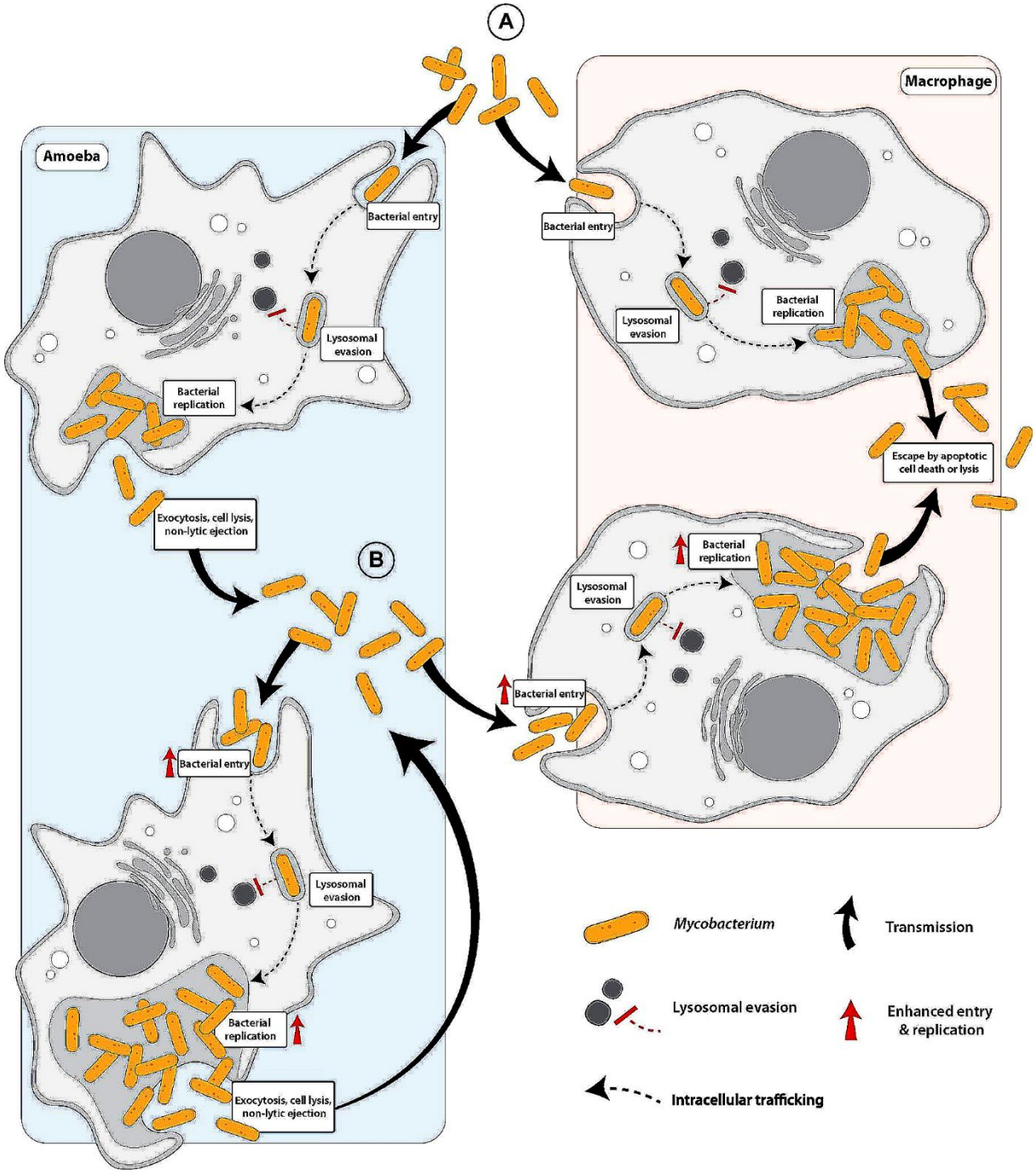

Figure 1.2. A proposed model for the adaptation of environmental mycobacteria to intracellular life. The model illustrates a potential evolutionary pathway in which ancestral *Mycobacterium tuberculosis* complex (*MTBC*) bacteria adapted to mammalian macrophages by first interacting with free-living amoebae. (A) Environmental mycobacteria are phagocytosed by amoebae but evade degradation by inhibiting phagolysosomal maturation. (B) These pre-adapted bacteria can subsequently infect macrophages, employing conserved evasion mechanisms to survive within this new, yet ecologically similar, intracellular niche (Price *et al*., 2024).

The biphasic evolutionary model of the *MTBC* underscores a significant and growing threat to global health security. This threat is two-fold: first, from the continual emergence of drug-resistant strains that undermine current treatments (Van Der Laan *et al*., 2022), and second, from the potential that environmental NTMs could undergo a similar evolutionary trajectory,

giving rise to new obligate human pathogens (Price *et al*., 2024). The risk is particularly heightened for opportunistic NTM pathogens such as *Mas*, where the evolution of a more efficient intracellular lifestyle—enabling optimized resource scavenging and transmission—could give rise to deadly pathogenic strains possessing inherent broad-spectrum resistance. Consequently, a strategic shift is needed to develop novel anti-mycobacterial therapies that circumvent the limitations of classical antibiotic drugs by exploiting new avenues, thereby remaining effective against drug-resistant strains.

1.4 Novel drug development strategies to combat existing and emerging drug-resistant mycobacterial strains

Significant research efforts are now focused on developing antibacterial agents that disrupt native prokaryotic protein homeostasis, with the caseinolytic protease (Clp) system emerging as a primary target. As an essential and nearly universal bacterial system, Clp has garnered considerable interest for the development of new drugs. A major obstacle in this endeavour, however, is the potential for off-target effects against homologous human systems.

The Clp protease system, particularly its ClpP core, is highly conserved in bacteria and in human mitochondria (Mabanglo and Houry, 2022). This homology creates substantial risks for off-target effects that could lead to mitochondrial dysfunction and damage to beneficial microbiome symbionts. The PPS is another native protein turnover pathway in prokaryotes. In contrast to the ubiquitous Clp systems, the PPS is largely restricted to actinobacteria, including mycobacteria, with only sporadic occurrences reported in a few other phyla such as Deltaproteobacteria and Planctomycetes (Iyer *et al., 2008*). Like eukaryotic ubiquitination, the PPS is a post-translational regulatory pathway that directs proteins for degradation. This is a rare mechanism in prokaryotes, employing a small protein tag (prokaryotic ubiquitin-like protein; Pup^{E}) to mark targets for degradation via a compartmentalized protease, the proteasome. Targeting the PPS presents a highly selective strategy for anti-mycobacterial drug development, as it is largely absent in other bacteria except for a few beneficial symbionts. Furthermore, this approach may pose a lower risk of resistance because pupylation is not essential for core proliferation, thereby reducing the selection pressure for escape mutations.

1.4.1 The mycobacterial PPS system

The PPS mechanism involves two stages: pupylation, the ATP-dependent ligation of a deamidated Pup^E tag to a target protein substrate, and proteasomal degradation, where the mycobacterial ATPase (Mpa) proteasome complex recruits the tagged protein for degradation (Von Rosen *et al*., 2021). The PPS is initiated when the deamidase of Pup (Dop) prepares Pup for the pupylation reaction by converting Pup^Q (C-terminal glutamine residue) to Pup^E (C-terminal glutamate residue) via a deamidation reaction. The prepared Pup^E tag is then bound and activated for ligation by the PPS ligase PafA via a phosphorylation reaction, which in turn catalyses the ligation of the now-activated Pup^E-P intermediate to a protein target's lysine residue, completing the pupylation reaction. Pupylated proteins are recognized by the 20S Mpa proteasomal complex for complete degradation (Figure 1.3).

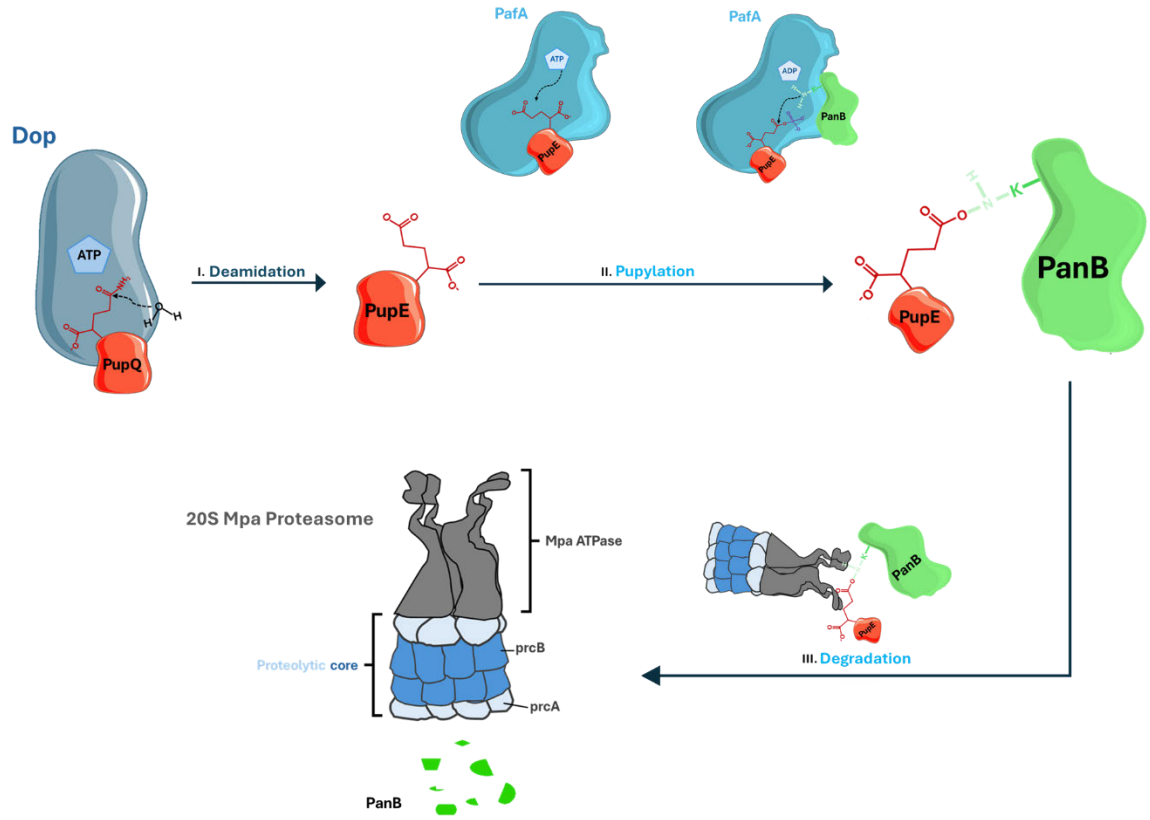

Figure 1.3. Schematic representation of the Pup-proteasomal degradation cascade system (PPS) in mycobacteria. Model protein target PanB proteasomal degradation via the PPS system is illustrated involving three key processes of the system. Initially, Pup^Q is (I) deamidated by the deamidase of pup (Dop) into Pup^E, which is subsequently bound and activated via phosphorylation by the proteasome accessory factor A (PafA). Once activated, the Pup^E-P intermediate is then ligated to the bound target PanB lysine residue by PafA, completing the (II) pupylation reaction. The pupylated PanB is marked for proteasomal degradation by the 20S Mpa proteasome complex. The Mpa recognizes PanB-Pup^E via interaction with Pup^E's N-terminal domain, inducing PanB unfolding and subsequent direction into the prcAB proteolytic core for degradative processing.

Not all pupylated substrates are fated for proteasomal degradation. Dop additionally functions as a depupylase, rescuing proteins in some cases from degradation by facilitating Pup^E removal. This depupylation activity is also thought to serve as a recycling mechanism, removing Pup^E from unfolded substrates prior to degradation and recycling the tag for the pupylation of other target proteins. Intriguingly, both Dop and PafA exhibit depupylation activity, but with a key distinction.

While Dop acts as a dedicated depupylase, Zhang *et al.* (2017) demonstrated that PafA can also catalyze a 'transpupylation' reaction—transferring Pup^E from a pupylated donor protein Ino1 to acceptor substrates FabD and Log (Zhang *et al.*, 2017). This functional overlap in activities by both enzymes is likely due to their significant structural homology (Figure 1.4), particularly in their active sites which similarly bind Pup and nucleotides. Their functional distinction arises from nucleotide specificity: PafA being an ATP-dependent ligase that phosphorylates its substrates Pup^E directly (Striebel *et al*., 2009; Burns *et al*., 2010; Özcelik *et al*., 2012), whereas Dop binds and hydrolyzes ATP to ADP and inorganic phosphate which remain bound for its deamidation/depupylation functions. Surprisingly, however, was the observed superior depupylation efficiency of PafA compared to the dedicated depupylase Dop, a phenomenon that can possibly be explained by PafA's higher affinity to bind pupylated substrates, binding pupylated substrates more effectively to catalyze Pup^E removal compared to Dop, which primarily binds PupE (PafA having multiple sites for both substrate recognition and Pup^E binding for the ligation reaction) (Zhang *et al.*, 2017).

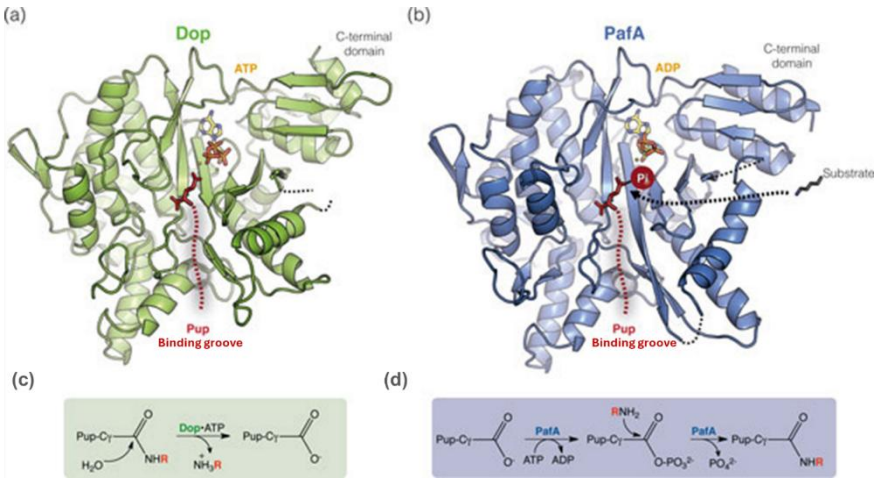

Figure 1.4. **Dop and PafA structural homology.** Dop (a, green) and PafA (b, blue) structural similarities illustrating the pup-binding groove and nucleotide-binding site. The PafA structure illustrates a bound phosphorylated Pup^E-p intermediate, activated for target protein substrate ligation upon substrate binding onto the substrate-binding site (indicated by an arrow). Dop catalyzes the deamidation of Pup^Q C-terminal glutamine residue into Pup^E (c), which is subsequently phosphorylated and ligated to protein substrates by PafA (d) (Barandun *et al.*, 2012).

Unlike eukaryotic ubiquitination, the pupylation reaction is solely mediated by the ATPase PafA ligase, which activates Pup^E via the phosphorylation of the γ-carbon of its C-terminal glutamate residue and subsequently transfers the PupE-p intermediate to the target protein's lysine residue (Striebel *et al*., 2010). The Pup^E-protein complex is then recognized by the 20S Mpa proteasome, an intricate heteroprotein complex formed by the association of the Mpa ATPase and the protease cores, specifically the proteasome subunits alpha and beta (prcAB), which form an efficient compartmentalized protease system for degradation, the product of PPS (Figure 1.3). Mpa recognizes the pupylated protein complex by interacting with ligated Pup^E's N-terminal coiled-coil domain, triggering ATP-dependent unfolding of the entire pupylated substrate, and the translocation of the unfolded substrate towards the prcAB proteolytic cores for degradation. Importantly, the 20S proteasome's ability to distinguish between free Pup^E and Pup^E-substrate complexes stems from a critical conformational change in Pup^E. Upon ligation to a protein substrate, Pup^E undergoes a key conformational shift: its intrinsically disordered N-terminus folds into an ordered coiled-coil domain. This structured form is specifically recognized by Mpa, directing the complex to the 20S proteasome for degradation.

1.4.2 PPS role in stress response

Mycobacteria encounter various stresses depending on their habitat, from the limiting environment for most NTM to intracellular microbicidal agents for the pathogenic species like the *MTBC*. In response, they have evolved several mechanisms to aid survival and proliferation. These adaptations have produced durable populations of species capable not only of enduring and proliferating under adverse conditions, but also robust in responding and adapting to new stresses (Yimcharoen *et al*., 2023). Mycobacterial species, particularly in the pathogenic classes like the *MTBC,* demonstrate robust responses to stressful conditions, rapidly responding to numerous host-derived antimicrobial stresses by inducing protective metabolic cascades.

The stringent response (SR) in *MTBC* is the most well-known example, playing a pivotal role in virulence and resistance via bacterial persistence in the nutrient-limited macrophage. SR is a conserved bacterial adaptive response system that induces metabolic dormancy to ensure long-term survival during nutrient deprivation. This response is triggered by the accumulation of the alarmone guanosine tetraphosphate and pentaphosphate (collectively (p)ppGpp) and inorganic polyphosphate, which in turn suppresses growth-related metabolic processes (Danchik *et al*.,

2021). Elevated (p)ppGpp levels initiate a signalling cascade that globally reprograms gene expression, downregulating protein synthesis machinery while upregulating stress response genes. This reprogramming forms a crucial strategy in *MTBC* persistence and dissemination, ensuring long-term survival in host cells by establishing latent TB infection, which can cause active disease and spread to new hosts. This latency constitutes the largest reservoir of the pathogen. As these infections are typically asymptomatic until reactivation, SR enables clinical evasion and perpetuates the *MTBC* population, making eradication exceedingly difficult.

A major consequence of this facilitated dormant state is antibiotic tolerance, a key mechanism of persistence. SR indirectly contributes to drug resistance by reducing metabolic activity, thereby making the pathogens less susceptible to antibiotics that target active metabolic activity (Ahmad *et al*., 2025). This simple yet effective strategy is central to *MTBC's* ability to endure adversity through the global change in gene expression by (p)ppGpp under stress, which is linked to the upregulation of stress response genes, including those involved in protein turnover systems such as the Clp protease system and the PPS in mycobacteria. The PPS, in particular, is heavily implicated in mycobacterial stress response, playing a significant role in the pathogen's persistence under the hostile environment of the host cells.

1.4.3 PPS role in aiding mycobacterial pathogenicity

The PPS is a well-established virulence factor and a promising target for novel TB therapies. Since the systems discovery and outline by Weber-Ban's research team (Weber-Ban *et al.*, 2012), extensive work by numerous groups has explored delineating the entire system, from the pupylation ligation reaction that tags protein targets with Pup^{E} to the degradation of the tagged proteins by the 20S Mpa proteasome. However, the regulatory mechanisms controlling the PPS—whether at the genetic or enzymatic level—remain poorly understood, though it is known that stress conditions encountered during primary *MTBC* infection significantly upregulate the system. A longstanding presumption was that the PPS is upregulated alongside the Clp protease system as part of the global transcriptional reprogramming during the stringent response.

Curiously, PPS-related genes cluster with either each other or other stress-related genes on the same locus (Figure 1.5), naturally implying an operon-controlled system towards a regulated stress response. This has, in part, been confirmed by Festa and coworkers, who characterized the presence of a proteasome accessory factor (Paf) operon in *Mtb,* an operon that includes the

pupylation integral protein Pup^{E}-ligase "PafA" (Festa *et al.*, 2007). Using PCR, the group demonstrated that *Mtb* PPS PafA is co-transcribed with two stress-related genes i.e., PafA, PafB, resulting in the so-called "PafABC operon".

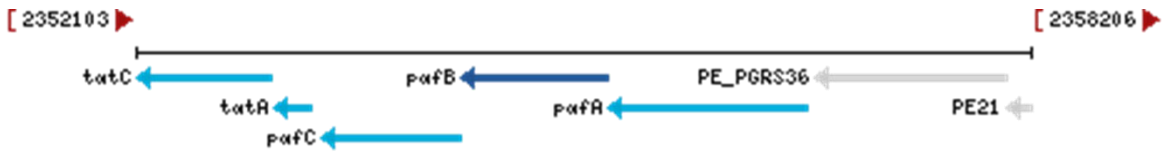

Figure 1.5 The PafABC operon in *Mycobacterium tuberculosis* (*Mtb*). architecture of the *pafABC*(+) operon in *Mtb* H37Rv (coordinates 2 352 103 - 2 358 206) housing the core proteasome accessory factor (Paf) operon genes PafA, PafB, and PafC, i.e., PafABC. Upstream of the operon sits two virulence-related PE/PPE family genes PE21 and PE_PGRS36, and downstream of the operon are additional virulence-related genes including the twin-arginine translocation (tat) pathway genes tatA and tatC (tatAC), the RNA-activated DNA helicase (hel) gene helY, and the conserved uncharacterized gene Rv2091c. All the genes are directional with the PafABC operon, clustering together to imply a single long operon inclusive of PafABC i.e., the PafABC(+) operon transcribing upstream PE21 to Rv2091c.

Despite limited initial data on PafA's function, the study successfully established it as an indispensable component of the proteasome system (PPS). The authors used a genetic complementation approach to demonstrate this: transposon-disrupted *pafA* mutants exhibited dramatic accumulation of the PPS substrate Mpa, but this phenotype was reversed when the mutants were restored with either *pafA* alone or the entire *pafABC* operon (Figure 1.6). This PPS deficiency was confirmed by the accumulation of other established PPS targets, FabD and PanB, in the *pafA* mutants. Crucially, this effect was unique to PafA, as *pafB* and *pafC* mutants showed wild-type levels of Mpa, FabD, and PanB (the recombinant latter two degraded back to wild-type levels in these mutants), indicating that PafA is the primary enzymatic unit of the PPS within the operon (Festa *et al.*, 2007).

The functional analysis of the operon was extended to a physiological context, revealing that the entire *pafABC* operon is essential for full resistance to reactive nitrogen intermediates (RNI)—key antibacterial agents produced by host immune cells. Upon RNI exposure, bacterial growth was severely impaired only in the *pafA* mutants. In contrast, *pafB* and *pafC* mutants showed only slight growth defects compared to wild-type *Mtb*, suggesting that these genes play a minor, supportive role in RNI resistance (Festa *et al.*, 2007).

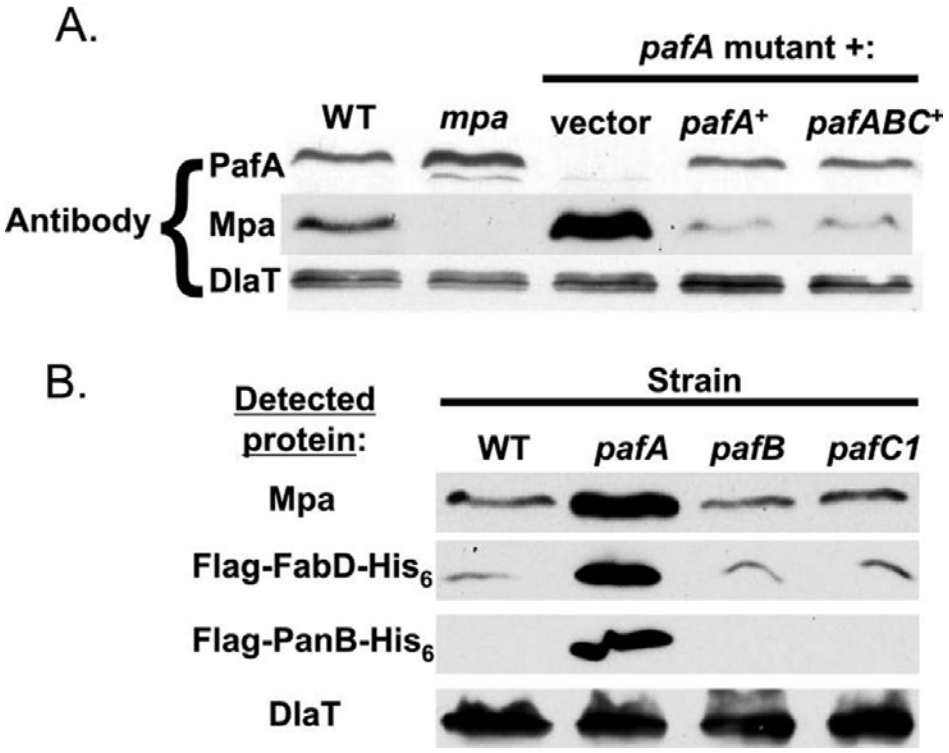

Figure 1.6 PafABC operon role in the pup proteasome degradation cascade system (PPS). PafA's exclusive role in the PPS was demonstrated via PafA mutagenesis and degradation studies. Using DlaT as a loading control, the figure demonstrates (A) PPS-target Mpa's significant increase in transposon-disrupted PafA *Mtb* mutants which is brought back to wild-type (WT) levels upon recombinant PafA expression in the same cell lines, and (B) illustrates the accumulation of PPS-targets Mpa, recombinant proteins FabD and PanB, in PafA transposon-disrupted *Mtb* mutants but not in transposon-disrupted PafB and PafC *Mtb* mutant cell lines, showing proteasome degradation of the PPS-targets back to WT levels only in the presence of PafA. The figure is extracted from Festa *et al.*, 2007.

It was not until a study conducted by the Weber-Ban group that the functional role of PafB and PafC in the RNI stress response was elucidated, linking both genes to a role in the DNA repair pathway (Fudrini Olivencia *et al.*, 2017). Their work revealed that PafBC knocked down *Mycobacterium smegmatis* (*Msm)* mutants exhibited significantly reduced levels of Recombinase A (RecA), a central component of the SOS-mediated DNA repair system. Consequently, these mutants exhibited markedly increased sensitivity to the established DNA-damaging agents UV irradiation and mitomycin C. This phenotype was rescued upon genetic restoration of *pafBC*, which restored survival to near wild-type levels. Furthermore, the study confirmed that, consistent with Festa *et al*. (2007), *pafB* and *pafC* are co-transcribed and that their protein products form a stable heterodimer.

Structural homology modelling revealed that the PafBC complex shares a high degree of similarity with known transcription regulators in actinobacteria and contains conserved DNA-binding motifs, suggesting its direct role in gene regulation. This hypothesis has since been addressed by the Weber-Ban group, who reported the intricate association of the PPS system with DNA damage response (DDR) (Schilling *et al.*, 2025). Through cryo-EM analysis, the group directly visualized how the PafBC heterocomplex is activated by the binding of short single-stranded DNA (ssDNA) within its WYL domain. This ssDNA binding triggers a critical interaction between PafBC and the DNA-dependent RNA Polymerase, reprogramming the enzyme to recognize a unique PafBC-specific -26 promoter element instead of the canonical -35 element (Figure 1.7).

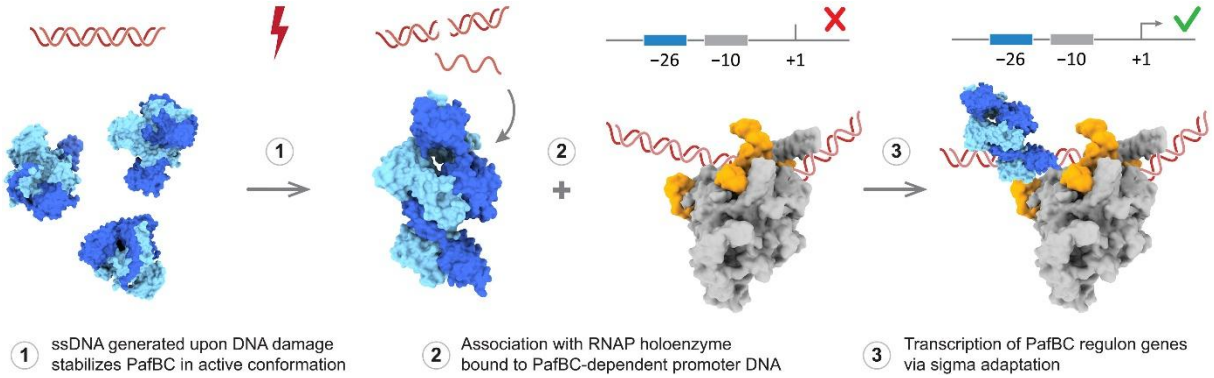

Figure 1.7. PafBC-mediated reprogramming of RNA polymerase in mycobacterial DNA damage response. The heteromeric PafBC complex is activated upon binding by single-stranded DNA (ssDNA) generated after DNA damage during stress (1). The activated complex then binds the RNA polymerase (RNAP) (2) and allosterically reprograms the enzyme to recognize a specific -26 promoter sequence, initiating the transcriptional response to DNA damage (Schilling *et al.*, 2025).

This promoter switch ultimately upregulates over 150 DDR genes, including key components of the parallel DDR SOS pathway (Figure 1.8). Notably, a majority of these PafBC-activated genes—including SOS response-related genes like *recA* and *sigG*—are part of the PPS regulon. Their protein products are established pupylation targets for the 20S Mpa proteasome, suggesting a feedback mechanism that degrades these stress-response factors once cellular homeostasis is restored (Figure 1.8).

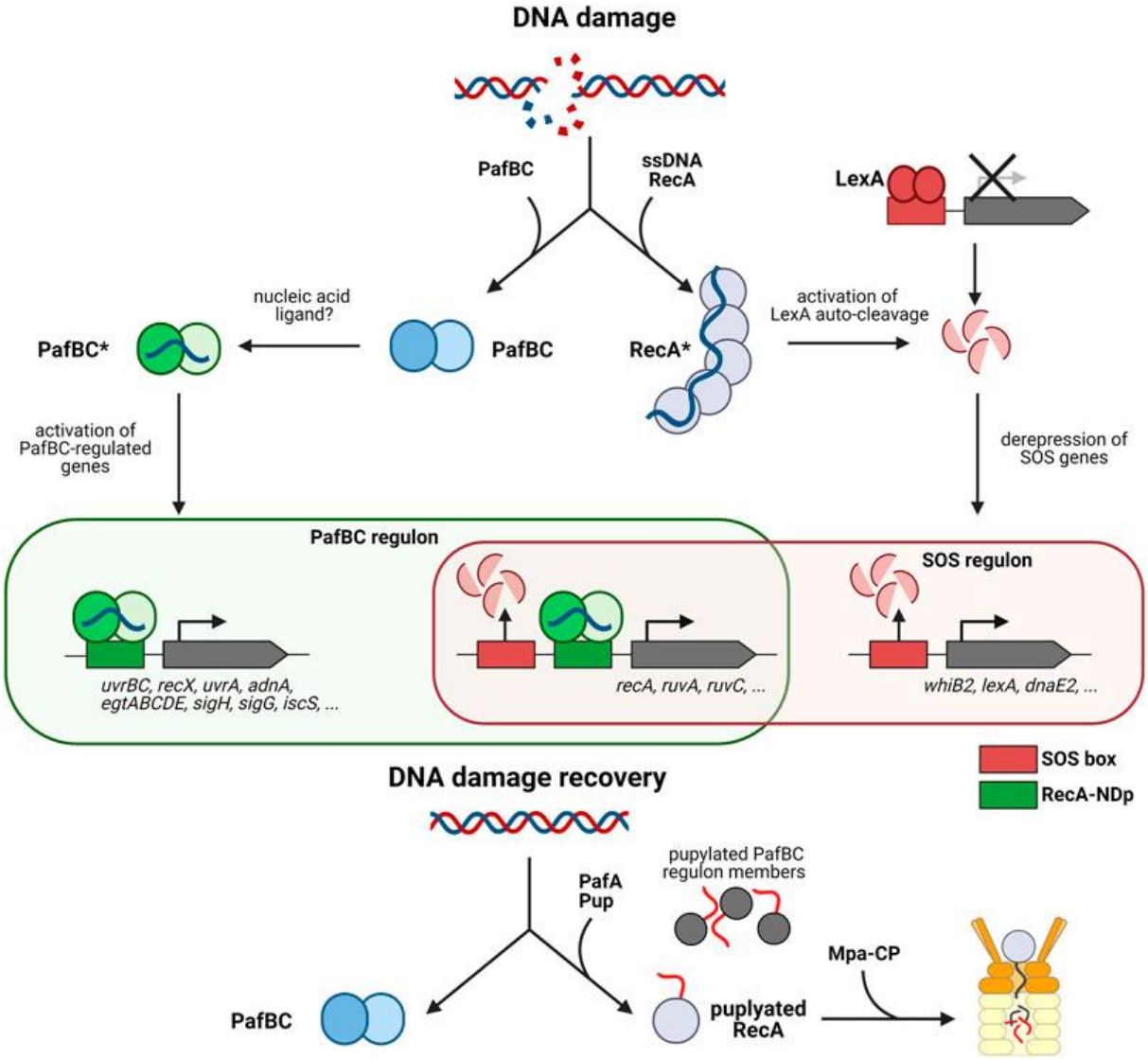

Figure 1.8. Coordination and regulation of DNA-damage repair by the PafBC and SOS pathways. DNA damage activates two primary response systems. The PafBC heterocomplex, activated by the binding of ssDNA, directly upregulates a specific regulon that includes DNA repair genes and *recA*, the master regulator of the SOS pathway. The induced RecA protein (also activated by ssDNA) then activates the canonical SOS response. These pathways function in tandem, with overlapping regulons (Venn diagram), to coordinate repair. During recovery, to restore homeostasis, key proteins from the PafBC regulon—including RecA—are tagged for degradation via the PPS and subsequently degraded by the Mpa-20S proteasome complex (Mpa-CP) (von Rosen *et al*., 2021).

Interestingly, genomic analysis of the PafABC operon locus reveals the presence of five additional stress-related proteins not included in the original Festa *et al*. (2007) study i.e., PE21 and PE_PGRS36 upstream of PafA, and tatA, tatC, helY, and the conserved Rv2091c downstream of PafC (Festa *et al.*, 2007). With the exception of Rv2091c—a predicted transmembrane protein that is non-essential for *in vitro* growth—every gene in this cluster contributes significantly to *Mtb* survival and pathogenesis. This gene association suggests a larger operon i.e., PafABC(+) operon, inclusive of the described PafABC operon and regulated

by a presumed native mycobacterial promoter upstream of PE21 gene and terminating downstream Rv2091c (Figure 1.5).

Partial support for this model comes from Festa *et al*., who demonstrated via transcriptional analysis that the *pafABC* operon is co-transcribed at low levels with the downstream gene *tatA*. The *tatA* gene is an essential virulence factor that forms part of the twin-arginine translocation (Tat) protein transport system, which facilitates the translocation of folded proteins across the cytoplasmic membrane to promote host interactions (Festa *et al.*, 2007; McDonough *et al.*, 2005). Importantly, however, was the lack of readthrough transcription into *tatC*. This suggested that *tatAC* is expressed from its own dedicated promoter, defining the downstream boundary of the proposed *pafABC*(+) operon and excluding *tatC* and subsequent downstream genes from its transcriptional unit. (Festa *et al.*, 2007). However, this observation may reflect a sophisticated regulatory mechanism. Rather than being constitutively transcribed in full, the *pafABC*(+) operon could undergo conditional expression, where only specific gene subsets are transcribed under basal conditions. The full operon may be transcribed only during defined stresses, such as the DDR when PafBC reprograms RNA polymerase. This would provide a more rapid and streamlined adaptation to host-elicited stress.

The remaining PPS-related genes are clustered at a separate locus upstream of the presumed *pafABC*(+) operon, suggesting their expression is coordinated by a distinct operon (Figure 1.9). This cluster includes, in sequential order, *dop*, pup^{Q}, *prcB*, and *prcA*. The *prcB* and *prcA* genes encode the core 20S proteasome subunits, which assemble with the Mpa ATPase to form the holoenzyme. These core subunits contain the proteolytic active sites responsible for digesting the unfolded protein substrates delivered by Mpa.

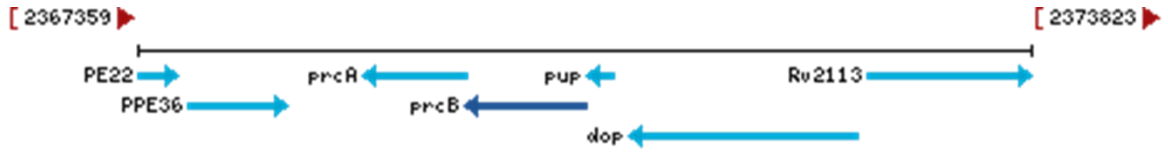

Figure 1.9 The hypothetical DopPupPcrAB operon in *Mtb*. *Mtb* (H37Rv) genome region 2367359 - 2373823 housing pup-protasome protein degradation system (PPS) genes dop (deamidase of pup), pup (prokaryotic ubiquitin-like protein), prcA and prcB (prcAB; proteasome subunit cores alpha and beta), encompassing the presumed DopPupPcrAB operon. Upstream the operon i.e., upstream dop is the Mtb-specific gene Rv2113, and downstream PE/PPE family genes PE22 and PPE36, respectively found opposite the direction of the presumed DopPupPcrAB operon.

The Mpa ATPase gene is also located near this presumed operon system, situated upstream Dop and separated by only the Rv2114 gene that encodes a hypothetical protein. The exclusion of Mpa ATPase from the presumed operon system could reflect the requirement of the ATPase in PPS independent processes, i.e., activity in the absence of the prcAB proteasomal core (as seen with other AAA+ family homologs mediating numerous cellular processes). This was partially supported by a mutagenesis study which demonstrated that Mpa retained its activity despite deletion of C-terminal amino acids that are now known to be essential for Mpa's association with the proteasome cores pcrAB (Darwin *et al.*, 2005).

The study additionally demonstrated that the deletion leads to a significantly increased sensitivity of *Mtb* to NTI stress, a hallmark reaction upon PPS disruption. This indicates that while ATPase activity is sufficient for the recognition and unfolding of pupylated substrates, the complete degradation of the protein by the proteasome cores *prcAB* is needed to confer RNI resistance. Furthermore, Mpa activity is regulated at the enzymatic level through post-translational modification. Independent studies have established that its ATPase function is controlled by the reversible pupylation of a specific C-terminal residue, lysine[591]. Monopupylation of this site by PafA deactivates Mpa, while depupylation by Dop restores its activity. This regulatory mechanism functions by disrupting the assembly of Mpa's conserved hexameric ring structure, which is essential for its function (Delley *et al.,* 2012; Festa *et al.*, 2010; Watrous *et al.*, 2010; Poulsen *et al.,* 2010). This aligns with findings from Darwin's group, which demonstrated that point mutations in the same conserved region also prevent hexamerization and deactivate Mpa. Therefore, both genetic mutation and post-translational pupylation achieve the same inhibitory outcome by physically impeding the formation of the functional hexameric complex

It is clear that the PPS system plays a significant role in mycobacterial regulated stress response, helping the pathogen evade the host's immune system and, consequently, significantly contributes to the pathogen's virulence. In the absence of the PPS, mycobacterial cells exhibit increased sensitivity to host immune responses, rendering them more susceptible to abrogation and removal by the host's natural defenses against infection over time. This susceptibility is demonstrated by independent *in vivo* studies using *Mtb*-infected mouse models. These studies highlight the critical role of the PPS system at all levels— from pupylation to Mpa-mediated 20S proteasomal degradation—during chronic infection. Pioneering work by groups independently led by Bishai and Ehrt collectively report *Mtb* attenuation in mice models

infected with proteasome-disrupted (Mpa ATPase or prcBA core, respectively) *Mtb* mutants (Lamichhane *et al.*, 2006; Gandotra *et al.*, 2007). This was illustrated by a significant decrease in *Mtb* burden in lung tissue post-exposure for the proteasome-disrupted mutants compared to the wild-type *Mtb*-infected mouse models. Between the two studies, mutant *Mtb* attenuation significantly increased mice survival by up to 121 days more than in wild-type *Mtb*-infected mice models, exposing the importance of PPS during chronic infection.

The PPS was demonstrated to be crucial for chronic infection; the mutants, though persisting, demonstrated comparatively slow growth that ultimately led to clearance in the mice's lung tissue. Notably, Lamichhane *et al*. (2006) study additionally reported that the mutant strains elicit a weak IFN-γ (an important cytokine that promotes a defensive immune response by activating macrophages during TB infection) host response, consequently limiting lung tissue damage which contrasts with the intense inflammation observed in wild-type *Mtb*-related infections. Interestingly, IFN-γ-silenced mouse models (models demonstrating a diminished ability to mount an effective anti-*Mtb* immune response) showed a significant increase in survival compared to the wild-type *Mtb*-infected models, suggesting that PPS may have a role beyond defense against IFN-γ-dependent immune responses. Importantly, Lamichhane *et al.* (2006) revealed that the PPS knocked-out mutant made for an efficient antigen for immunization because of the mutant's decreased virulence (persistence in diminished numbers and a weak IFN-γ response), resulting in a competent vaccine that provided protection from chronic infection by wild-type *Mtb*. PPS mutants (Mpa) were similarly attenuated in mouse-infected models in a study conducted by Darwin's group (Darwin *et al.*, 2003), additionally conveying the importance of the system for both *in vivo* and *in vitro* optimal growth where PPS mutants growth on agar was compromised, resulting in abnormal colonies.

1.4.4 PPS as an anti-mycobacterial target

PafA, the sole ligase of the PPS, presents a compelling drug target because its inhibition can fully disable the system. This disruption attenuates *Mtb* virulence, allowing the host's innate immune system to clear the infection in an IFN-γ-dependent manner. The PPS is particularly attractive as a target because it is not required for bacterial proliferation. This non-essential nature reduces the selection pressure for drug resistance, a major driver in the evolution of

resistant strains. Additionally, since the PPS contributes to antibiotic resistance, targeting the system could potentially rejuvenate the potency of existing antibiotics in combination therapies.

PafA's constitutive expression within the *PafABC* operon hints at potential allosteric regulation of the enzyme itself. This hypothesis is supported in a study by Ofer and collaborators. Through mathematical modelling of PafA enzymatic activity using engineered target protein substrates, the group predicted that *Msm* PafA is an allosterically modulated enzyme that cooperatively binds its substrates during pupylation (Ofer *et al.*, 2013). Their data revealed that *Msm* PafA binds protein targets and Pup^{E} independently, and that PafA allostery is exclusively triggered by substrate protein binding, a conclusion based on the finding that PafA remained unaffected in a Pup^{E}-binding assay (Ofer *et al.*, 2013). This observation, however, was made by assessing only PupE and the target protein as PafA ligands, with other components in the assay buffer (such as Mg^{2+} and ATP) also potentially playing a role in modulating PafA's allosteric behavior.

Recent work exploring the development of novel drugs capable of hijacking and repurposing another bacterial recycling cascade system i.e., the prokaryotic ClpP protease system, for targeted protein degradation has demonstrated encouraging results *in vivo*. Following the recent development and optimization of the groundbreaking class of molecules by Crews and collaborators, i.e., the protein targeting chimeras (PROTACs), recruitment of the versatile technology in other therapeutic ventures has been a subject of interest (Sakamoto *et al.*, 2001: Schneekloth *et al.*, 2008; Bondeson and Crews, 2017; Bondeson *et al.*, 2018; Pettersson and Crews, 2019; Pfaff *et al.*, 2019; Bond and Crews, 2021). Currently undergoing Phase II clinical trials and showing considerable promise, PROTACs are heterobifunctional proximity molecules that hijack the native protein turnover systems of cells to degrade a target protein (Figure 1.10). Possessing two functional molecules linked together via a flexible linker, PROTACs bind target molecules via a target protein ligand (warhead) and recruit the degradation system via the ligand on the opposite end (the recruiter), forming a ternary complex to facilitate a proximal tagging reaction that flags the protein for degradation (Liu *et al.*, 2020).

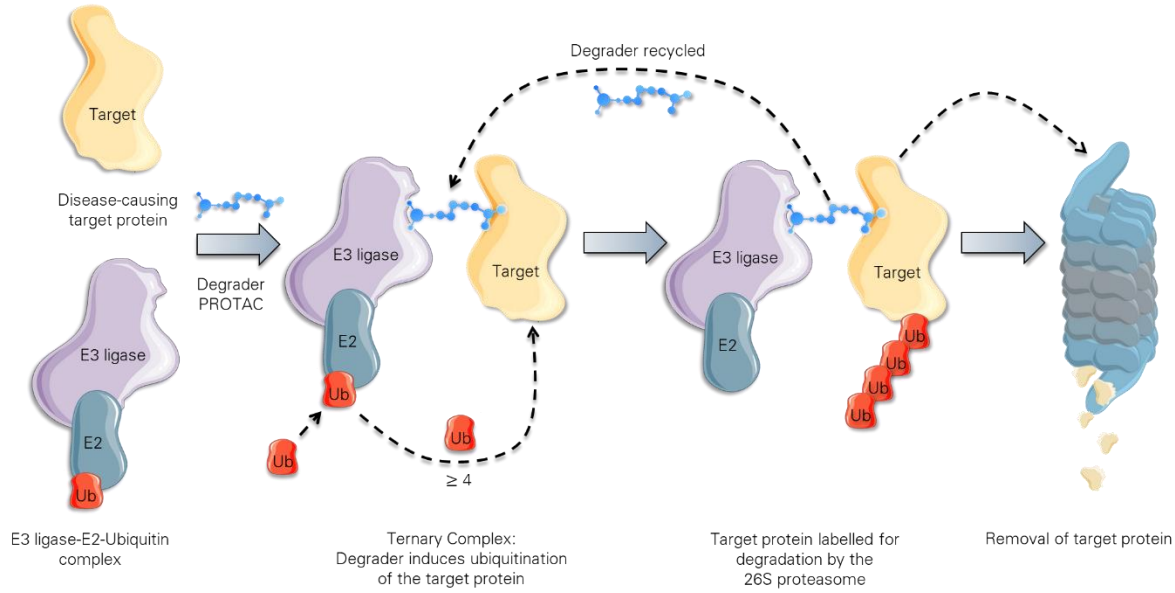

Figure 1.10 Schematic of PROTAC-mediated targeted protein degradation. (left-right) The bifunctional PROTAC molecule simultaneously engages a target protein via a warhead ligand and a specific E3 ubiquitin ligase via a recruiting ligand. This induces the formation of a ternary complex, bringing the target into proximity with the ubiquitination machinery (E3 ligase-E2-ubiquitin complex). The E3 ligase catalyses the polyubiquitination (typically ≥4 ubiquitin units) of the target protein. The polyubiquitinated target is recognized and degraded by the 26S proteasome. The PROTAC molecule is released and can catalytically facilitate multiple rounds of degradation (degrader recycling).

In addition to the achievements PROTAC technology has seen in cancer research, a team led by Clausen and Kaiser recently launched a study that applied PROTAC technology in a prokaryotic setting. The team developed three bacterial PROTAC-like molecules, namely BacPROTACs, that induced significant *in vitro* and *in vivo* protease-selective degradative activity of model protein substrates via the conserved ClpC:ClpP (ClpCP) system (Morreale *et al.*, 2022). To achieve this, the group created BacPROTAC molecules based on phosphorylated arginine residues (pArg) and cyclic antibiotic cyclomarin A (CymA), warheads that recruit the ClpCP bacterial protease.

In an *in vitro* study using *Bacillus subtilis*, one of the chimeric molecules BacPROTAC-1, created by linking the monomeric streptavidin (mSA) ligand biotin to pArg, led to the near-complete degradation of mSA at 1 µM (Morreale *et al.*, 2022). When the group conducted the same study in *Msm,* 1 µM BacPROTAC-1 hijacked the ClpCP mycobacterial homolog ClpC1P1P2 to cause almost complete selective degradation of mSA (Morreale *et al.*, 2022). To create mycobacterial selective BacPROTACs, the team additionally created BacPROTAC-2

and BacPROTAC-3 by linking CymA analog sCym-1 (a mycobacterial ClpC1P1P2 protease selective recruiter) to biotin and the bromodomain 1 (BRD1) ligand JQ1, respectively. At 1 μM, BacPROTAC-2 degradative activity analysis demonstrated significant mSA selective degradation comparable to pArg-based BacPROTAC-1 (Morreale *et al.*, 2022). Significant degradation was also observed for BacPROTAC-3, which selectively degraded the transcription regulator protein model BRD1, an established anti-cancer target (Fryland *et al*., 2016; Klein *et al*., 2018; Winter *et al*., 2015), at the same 1 μM concentration (Morreale *et al.*, 2022).

Another pioneering approach for targeted protein degradation in prokaryotes has successfully recruited the endogenous ClpXP protease. In a study parallel to that of Morreale *et al*. (2022), Izert-Nowakowska and colleagues developed a novel PROTAC-like molecule called CLIPPER (Clp-Interacting Peptidic Protein Eraser) (Izert-Nowakowska *et al*., 2025). The group designed a specific CLIPPER construct, a GroEL-Targeting Chaperonin (GroTAC), by fusing three elements: an SspB-derived "anchor" peptide that recruits *E. coli* ClpXP, a flexible linker, and the GroEL-binding SBP ligand. Notably, this GroTAC molecule was recombinantly expressed in *E. coli* using an arabinose-inducible system, enabling time-dependent targeted degradation of endogenous GroEL. Importantly, the successful progressive depletion of GroEL had profound downstream consequences, significantly disrupting the bacterial proteome and ultimately inhibiting cellular growth.

The successful development of BacPROTAC/CLIPPER molecules, which can selectively induce protein degradation by repurposing prokaryotic machinery, is promising and provides encouraging data for the development of more BacPROTAC drugs that can be applied in therapeutic interventions for diseases such as TB. Naturally, the mycobacterial PPS is an ideal foundation for novel BacPROTAC development. This approach would harness the PPS to degrade essential bacterial proteins via the 20S Mpa proteasome, leading to cell death. A key advantage is the potential to repurpose conventional antibiotics as warheads against drug-resistant strains. This is feasible because, as Bondeson *et al*. (2018) showed, PROTAC-induced degradation can occur even when the warhead has reduced binding affinity.

1.5 Study rationale

Given the significance of the PPS in mycobacterial pathogenesis—specifically, its role in sustaining the pathogen during the host's adverse immune response—targeting this cascade system for anti-proliferative therapy presents an intuitively promising approach. PafA, the pupylation integral protein singularly mediating Pup^E ligation onto protein substrates, presents an ideal target for the development of PPS targeting drug(s). Ideally, these novel molecules should possess innovative modalities that overcome challenges faced by current anti-TB therapeutics, including drug resistance.

This work aims to develop novel therapeutic strategies against *Mtb* by targeting its PPS, a bacterial degradation pathway analogous to eukaryotic ubiquitination. To this effect, focus was made on the central ligase, PafA, with the goal of discovering small-molecule ligands that can either inhibit the entire PPS or serve as recruiters for targeted protein degradation (BacPROTACs). Therefore, the studies specific primary objectives were to:

1. Biophysically characterize *Mtb* PafA.
2. Identify and validate novel small-molecule ligands of *Mtb* PafA
3. Evaluate the therapeutic potential of validated ligands by determining their effect on mycobacterial viability and stress response under PPS-pathway-inducing conditions

1.6 Novelty

This study presents a high degree of novelty by advancing the PPS of mycobacteria as a tractable and mechanistically validated drug target, with specific focus on the Pup ligase, *Mtb* PafA. Although the PPS has previously been recognised as essential for mycobacterial stress tolerance and virulence, this thesis moves beyond conceptual targeting to provide the first integrated biochemical, biophysical, and ligand-discovery framework demonstrating PafA ligandability and therapeutic potential. The work is particularly original in establishing that *Mtb* PafA exists as a constitutively active monomer regulated allosterically by Mg^{2+}—a mechanistic insight that challenges prior assumptions of oligomer-dependent activation in ubiquitin-like ligases. Furthermore, the study develops and optimises both TSA and CETSA platforms specifically for *Mtb* PafA, enabling the first experimental validation of small-molecule binding

to this PPS enzyme in both *in vitro* and cellular contexts. The identification of multiple structurally distinct ligands that stabilise PafA provides the first chemical starting points for PPS-directed anti-tubercular drug design. By integrating recombinant protein biophysics, structural modelling, molecular docking, and phenotypic assays, this thesis establishes PafA as a ligandable and druggable node within the PPS. Collectively, this work shifts the PPS field from descriptive biology toward actionable drug discovery, opening a new anti-mycobacterial strategy that is mechanistically distinct from classical antibiotic targets and therefore highly relevant for combating drug-resistant tuberculosis.

1.7 Significance and impact

This doctoral study makes a significant and timely contribution to anti-tubercular drug discovery by advancing the mycobacterial PPS ligase, PafA, as a novel and structurally tractable therapeutic target. Tuberculosis remains the leading cause of death from a single infectious agent globally, and the accelerating emergence of multidrug-resistant *Mtb* strains necessitates innovative strategies that move beyond classical antibiotic targets. By establishing PafA as a ligandable enzyme within the PPS—an actinobacteria-restricted protein degradation pathway essential for stress survival and virulence—this work identifies a highly selective vulnerability in mycobacteria with minimal risk of off-target effects in humans or commensal microbiota. The study's integrated experimental and computational approach, encompassing recombinant protein production, biophysical stabilisation assays (TSA/CETSA), and molecular docking, provides robust mechanistic evidence that small-molecule ligands can engage and modulate PafA function. Importantly, the identification of multiple validated ligands demonstrating significant thermal stabilisation establishes a credible foundation for rational PPS-directed inhibitor development. Beyond target validation, this work advances methodological platforms for screening and characterising PPS modulators, thereby enabling future structure-guided drug optimisation. Finally, the thesis addresses a critical gap in anti-mycobacterial therapeutics by expanding the druggable target landscape to include proteostasis pathways, offering a promising translational trajectory toward new agents active against drug-resistant tuberculosis.

CHAPTER 2. MATERIAL AND METHODS

The expression vector pET24b(+)PafA, encoding the pupylation ligase PafA, served as the foundational genetic construct for this study. This construct was obtained as a gift from the laboratory of Professor Heran Darwin (Department of Microbiology, New York University Grossman School of Medicine). The vector construct harbouring the *Mtb* (H37Rv) *pafA* gene was previously validated by Sanger sequencing and primer-specific PCR amplification, confirming 100% sequence identity (see Figure S5 in the Appendix).

2.1. Recombinant expression and purification of pupylation proteins

For recombinant expression of PafA, the pET24b(+)PafA vector was transformed into chemically competent BL21 (DE3) *E. coli* cells. Transformation was performed via a modified heat-shock method based on a standard calcium chloride ($CaCl_2$) protocol (Sambrook and Russell, 2001).

2.1.1 Competent cell preparation and transformation

Competent BL21(DE3) *E. coli* cells were prepared for transformation using a calcium chloride method. A single colony, obtained from an overnight streak of the non-recombinant cells on Luria-Bertani (LB) agar [10 g/L tryptone, 10 g/L NaCl, 5 g/L yeast extract, 15 g/L agar], was inoculated into 10 mL of LB broth. This starter culture was grown overnight at 37 °C with shaking at 200 rpm. Following incubation, the culture was diluted 1:100 into fresh LB broth and incubated under the same conditions until the optical density at 600 nm (OD_{600nm}) reached 0.4, as measured using a SpectraMax ABS Plus microplate reader (Molecular Devices, California, USA). The cells were harvested by centrifugation (5 000 × g, 15 min, 4 °C), gently resuspended in 40 mL of ice-cold sterile $CaCl_2$ solution (60 mM $CaCl_2$, 10 mM HEPES, pH 7.0), and recentrifuged. The final cell pellet was resuspended in 2 mL of the same ice-cold $CaCl_2$ solution to generate chemically competent cells, which were used immediately or aliquoted and stored at -80 °C.

For recombinant expression of PafA, chemically competent BL21 (DE3) *E. coli* cells were transformed with the pET24b(+)PafA vector. Approximately 90 ng of vector DNA was gently mixed with 20 µL of competent cells and incubated on ice for 30 min. A heat shock was applied by transferring the mixture to a 42 °C water bath for 90 s, followed by immediate incubation

on ice for 2 min. Following this, 80 µL of pre-warmed super-optimal catabolite medium [2% (w/v) tryptone, 0.5% (w/v) yeast extract, 10 mM NaCl, 10 mM $MgCl_2$, 20 mM glucose, 2.5 mM KCl, and 10 mM $MgSO_4$] was added to the now transformed cells, and the culture was recovered for 1 h at 37 °C with gentle shaking. Subsequently, 50 µL aliquots of the transformation mixture were spread onto pre-warmed LB agar plates supplemented with kanamycin (100 µg/mL) and incubated overnight statically at 37 °C. To confirm successful transformation, individual colonies were randomly selected, inoculated into selective LB broth for overnight culture (37 °C, 200 rpm), and each respective culture was subjected to plasmid isolation followed by restriction digestion and PCR analysis. Positive clones were cultured and archived as 25% (v/v) glycerol stocks stored at -80°C.

2.1.2 GroEL-GroES assisted recombinant co-expression of PafA

To promote the soluble, functional expression of recombinant PafA, a co-expression strategy with the chaperone system GroEL/GroES was employed. For this purpose, the pGro7 vector (Takara Bio Inc., Kusatsu, Japan), encoding the *groEL* and *groES* genes, was introduced into the previously generated recombinant pET24b(+)PafA BL21(DE3) *E. coli* strain (see Section 2.1.1). The transformation protocol followed the method described in *Section 2.1.1*, with the modification that all subsequent growth media were supplemented with kanamycin (100 µg/mL) and later both kanamycin (100 µg/mL) and chloramphenicol (25 µg/mL) to maintain the pET24b(+)PafA and pGro7 vectors following their transformation.

A single colony of the confirmed recombinant pET24b(+)PafA-pGro7 BL21(DE3) *E. coli* strain was obtained by three-way streaking from a glycerol stock onto a double-selective LB agar plate (50 µg/mL kanamycin, 25 µg/mL chloramphenicol), which was incubated overnight at 37 °C. An isolated colony was used to inoculate 10 mL of selective Terrific Broth (TB) medium [1.6% (w/v) tryptone, 1.0% (w/v) yeast extract, 0.5% (w/v) NaCl, and potassium phosphate buffer made up of 0.23% (w/v) KH_2PO_4 and 1.25% (w/v) K_2HPO_4 - 0.45 µM filter sterilized into autoclaved medium], supplemented with the same antibiotics, and grown overnight (37 °C, 200 rpm). This starter culture was then diluted 1:100 into a larger volume of fresh TB containing antibiotics and 0.5 mg/mL L-arabinose to induce chaperone expression. The culture was grown at 37 °C with shaking (150 rpm) until OD_{600nm} reached 0.8–1. To facilitate soluble *Mtb* PafA recombinant expression, the temperature was lowered to 25 °C and incubation continued for 24 hours with shaking. Following incubation, the bacterial cells were

harvested by centrifugation (5,000 × g, 15 min), and the pellet was either processed immediately for protein purification or stored at –20 °C.

2.1.3 Recombinant PafA purification

A sequential two-step purification procedure was employed to obtain pure recombinant PafA including immobilized metal affinity chromatography (IMAC) and size exclusion chromatography (SEC)

IMAC purification of recombinantly expressed PafA was done using Ni^{2+}-NTA resin (Thermo Fischer Scientific, Massachusetts, USA) following the manufacturer's instructions, with amendments. Initially, pellet culture of recombinantly expressed PafA was resuspended fully in sodium phosphate equilibration buffer (1xEB; 2.02% (w/v) $Na_2HPO_4.7H_2O$, 0.34% (w/v), NaH_2PO_4, 10% (v/v) glycerol, 300 mM NaCl, pH7.4) supplemented with 1 mM Phenylmethylsulfonyl fluoride Protease Inhibitor (Thermo Fischer Scientific, Massachusetts, USA), sonicated for 20 min (at 60% power, 30 s on 30 s off) using the BIOBASE Intelligent Ultrasonic Processor (Shanghai Xiwen Biotech, Shanghai, China), and the solution centrifuged at 10 000 x *g*, 4 ℃ for 30 min.

The supernatant fraction was filtered through Whatman No.1 filter paper, gently added to a prepared Ni^{2+}-NTA resin preincubated for 30 min in 1xEB buffer, and the solution subsequently incubated at 4 ℃ overnight with end-over-end rotary shaking. After incubation, the resin solution was allowed to rest by gravity flow (~ 10 min), and the unbound soluble fraction was removed by pipetting. The resin was sequentially washed with 50 mL 1xEB to further remove unbound proteins (W1), washed twice for 1 h end-over-end rotary shaking with wash buffer (1xEB supplemented with 50 mM imidazole, pH 7.4), and washed overnight end-over-end rotary shaking at 4 ℃ with wash buffer. Post washing, the protein was eluted with elution buffer (1xEB supplemented with 250 mM Imidazole, pH 7.4), and the elutions samples collected into sterile tubes containing concentrated dithiothreitol (DTT) to make a final 1 mM DTT solution.

SEC purification of IMAC purified PafA was done using the GE ÄKTA Prime Plus FPLC System (Cytiva, Marlborough, USA) equipped with a Sephacryl S-200 resin (Cytiva, Marlborough, USA) for resolution and PrimeView 5.31 (GE Healthcare Technologies Inc., Chicago, USA) software for analysis. Following the manufactures instructions the ÄKTA system was washed with filtered 1xEB buffer (supplemented with 1 mM DTT, pH 7.4), adapted

with the Sephacryl S-200 resin, the column respectively washed and equilibrated with 60 mL filtered deionised H_2O and 120 mL 1xEB buffer (supplemented with 1 mM DTT, pH 7.4) at 1 mL/min, 2.5 mL of the prepared Ni^{2+}-NTA purified PafA sample (centrifuged at 20 000 x *g* for 10 min, 4 °C) injected onto column, and sample separation and elution carried out with 120 mL 1xEB buffer (supplemented with 1 mM DTT) applied at 0.5 mL/min flow rate collecting 2 mL elution fractions. Using the UV spectrum generated via PrimeView, fractions were selected for SDS-PAGE and immunoblot analysis to assess PafA purification quality. Using the generated elution curve, selected samples were pooled and concentrated using Amicon Ultra centrifugal filters (Merck, Darmstadt, Germany) with a 10 kDa molecular weight cut-off.

2.1.4 Polyacrylamide gel electrophoresis and western blot analysis

2.1.4.1 Denaturing and native PAGE analysis

For SDS-PAGE analysis, protein samples were mixed 1:1 (v/v) with 2 x Laemmli sample buffers (reducing or non-reducing), and solutions were denatured at 95°C for 5 minutes prior to loading. Insoluble fractions obtained from cell lysate pellets were resuspended in 1xEB to match the volume of the corresponding soluble fraction, and the solutions similarly mixed 1:1 with 2 x Laemmli buffer. All samples were loaded onto prepared 10% polyacrylamide gels alongside the Spectra Multicolour Broad Range Protein Ladder (Thermo Fischer Scientific, Massachusetts, USA, Cat. No. 26619). Electrophoresis was performed in Tris-glycine-SDS running buffer at a constant 90 V until the dye front reached the bottom of the gel, and the resolved proteins were visualized by staining with Coomassie Brilliant Blue R-250.

Native PAGE was performed to analyse proteins in their non-denatured state. Gels (10% polyacrylamide) and Tris-glycine running buffer were prepared identically to those used for SDS-PAGE, but without the inclusion of denaturing agents, including SDS and reducing agents. Protein samples were prepared in native sample buffer [62.5 mM Tris-HCl (pH 6.8), 10% glycerol, 0.001% bromophenol blue] and loaded without prior heating. Electrophoresis was carried out at a constant 90 V until the dye front reached the bottom of the gel, and the resolved proteins were visualized by staining with Coomassie Brilliant Blue R-250.

2.1.4.2 Western blot analysis

For immunoblot analysis, proteins from unstained gels (including native PAGE gels) were transferred onto 0.2 µM nitrocellulose membranes (Cytiva, Marlborough, USA) using a semi-dry transfer system (Thermo Fischer Scientific, Massachusetts, USA) at a constant 10 V for 1 h at room temperature (RT). Following transfer, membranes were blocked for 1 h at RT in 5% (w/v) non-fat Blotto milk powder (Santa Cruz Biotechnology, Texas, USA) prepared in Tris-buffered saline (TBS; 20 mM Tris, 150 mM NaCl, 0.001% Tween-20, pH 7.4). Membranes were then washed 3 x for 5 min with TBS and subsequently incubated with mouse monoclonal anti-6×His primary antibody (Santa Cruz Biotechnology, Texas, USA, Cat. No.: SC-53073) diluted 1:5 000 in 0.5% (w/v) bovine serum albumin (BSA) TBS overnight at 4°C, gentle agitation. following 3 x 5 min washes with TBS, membranes were incubated for 2 hours at RT with a horseradish peroxidase (HRP)-conjugated goat anti-mouse secondary antibody (Thermo Fischer Scientific, Massachusetts, USA, Cat. No.: R-05074-500) diluted 1:10,000 in 0.5% (w/v) BSA-TBS. Following a final wash series (3 × 5 min with TBS), immunoreactive bands were detected using enhanced chemiluminescence (ECL) substrate (Bio-Rad, California, USA, Cat. No.: 170-5060) according to the manufacturer's instructions. Chemiluminescence signals were captured using a G:BOX Chemi XR5 imaging system (Syngene) and analyzed with GeneTools software (version 1.8.0).

2.2 PafA ligandability studies

2.2.1 PafA TSA optimization

A PafA thermal shift assay was designed and optimised by varying purified recombinant PafA concentrations (0.2; 0.5; 0.6; 0.8; and 1 µg) against 2X and 10X SYPRO Orange dye concentrations. Each assay was done by sequentially preparing the protein-buffer mixture, incubating the solution for 1 h at RT, followed by the subsequent addition of SYPRO Orange dye stock solution (made in 1xEB buffer supplemented with 1 mm DTT, pH 7.4) and the solution kept in the dark until thermal melt analysis. Fluorescence analysis was done using the precision QuantStudio5 Real-Time PCR instrument (Thermo Fischer Scientific, Massachusetts, USA) starting thermal melt reactions from 25 - 99 °C at 0.05 °C/s increment standard mode. The instrument analysis settings included ROX dye as the reporter, analysing fluorescence data

at 470 nm ± 15 excitation and 586 nm ± 10 emission wavelengths. Data generated after the assay was analysed using the Protein Thermal Shift Software version 1.4 (Thermo Fischer Scientific, Waltham, USA). Derivative thermal melt data from the thermal shift software was subsequently exported and thermal melt derivative curves generated using GraphPad Prism 9.2.0.332 software.

2.2.2 PafA TSA: Small-molecule ligand screening

Thermal melt assays to assess compound binding to PafA were performed using the optimised PafA thermal melt assay conditions described previously. Briefly, 1 µg of purified recombinant PafA was incubated with 10 µM of the respective ChemBridge DiverSET compound in 1× EB buffer for 30 minutes at room temperature. Following incubation, SYPRO Orange dye was added to a final concentration of 10×, and thermal melt analysis was conducted (see Table 2.3). The compound concentration (prepared in DMSO) was adopted from the previously optimised non-specific aggregator assay, which was used to screen compounds for non-specific aggregator properties. This concentration thus served as a primary basis for assessing binding under conditions where the compounds are known not to cause aggregation-induced activity. Control ligands for PafA were prepared similarly for TSA analysis to verify assay veracity. The bona fide nucleotide ligand ATP and two potent described PafA inhibitors—bithionol and ST1926 (binding the conserved ATP pocket and pup-binding groove of mutated *Mtb* PafA, respectively)—were selected. ATP and Mg^{2+} were used at concentrations equivalent to those in the in vitro pupylation assay (2.5 mM). Because $MgCl_2$ concentrations exceeding 7 mM interfere with TSA, PafA was pre-incubated with 20 mM $MgCl_2$ (equivalent to in vitro pupylation conditions) prior to use in TSA buffer containing 5 mM $MgCl_2$. The apo protein control was prepared using TSA buffer without $MgCl_2$ additions.

Table 2.1. Preparation of PafA TSA reactions

Reagent	Stock concentration	Volume (µL)	Final concentration
1xEB Buffer	-	15.3	-
PafA*	500 ng/µL	2	50 ng/µL
Ligand	1 mM	0.2	10 µM
Dye*	80 x	2.5	10 x

* Stocks made in assay buffer 1xEB

2.2.3 PafA CETSA: *in vivo* PafA ligandability studies

Compounds shown to bind PafA in the PafA TSA were subsequently further assessed for PafA binding in a cellular environment using recombinant BL21 (DE3) *E. coli* cells. PafA ligandability was assessed on recombinantly expressed PafA.

PafA recombinant expression in BL21 (DE3) *E. coli* cells was facilitated as detailed previously in 400 mL via coexpression with the GroEL-GroES chaperone system by autoinduction. following recombinant PafA expression, cell culture was harvested by centrifugation (5000 x *g*, 15 min), resuspended in 15 mL 1xEB buffer, sonicated for 10 min (60% power, 30 s on, 30 s off), and the soluble fraction subsequently recovered after centrifugation at 10 000 x *g*, 30 min. The supernatant was separated as 1.2 mL aliquots in autoclaved 1.5 mL polystyrene microtubes and subsequently supplemented with the respective compounds at 10 µM final concentration, incubated overnight at 4 °C with gentle end-over-end rotary agitation. Post compound exposure, each respective sample was separated into autoclaved PCR tubes as 100 µL sample fractions, and individual fractions incubated at assigned respective temperatures i.e., temperatures lower and higher than the elucidated PafA T_m (37, 40, 45, 47, 50, 54, 59, 65, 75, 80, and 85 °C) for 6 min in the PCR instrument (Applied Biosystems, Massachusetts, USA). Two control samples were included in every run, each respectively incubated at RT and on ice.

Post exposure, the samples were then allowed to cool at RT for 6 min, and samples centrifuged at 20 000 x *g* for 30 min at 4 °C. For each sample, supernatants from the respective fractions were prepared and separated on a 10.5% w/v reducing SDS-PAGE gel along with a PageRuler Prestained Protein Ladder (Thermo Fischer Scientific, Massachusetts, USA), semidry blot transferred (Semi-dry Western Blotting Apparatus, Consort, Lucknow, India) for 1 h at a constant 20 V onto a 0.2 µM nitrocellulose membrane, and the transfer efficiency analyzed by Ponceau staining. Transferred recombinant PafA was immunodetected using primary mouse monoclonal anti-His and secondary goat anti-mouse monoclonal antibodies as detailed previously. PafA detection was facilitated by clarity ECL substrate (Cytiva, Marlborough, USA) and imaging done using the G:BOX Chemi XR5 imaging system (Syngene, Karnataka, India) with the installed GeneSys software version 1.8.0.

For each experimental transfer, monomeric PafA bands were subjected to densitometry analysis using the Image J software. Relative intensity values for each experimental sample were calculated against the control samples, with the highest intensity samples (PafA incubated at

RT) assigned an intensity level of 100%, and the lowest intensity sample (a no-sample control, SDS-PAGE sample treatment buffer) assigned an intensity level of 0%. Data from the software was analyzed using GraphPad Prism 9.2.0.332 software (GraphPad Prism, San Diego, USA), normalized and used to generate non-linear regression models used to calculate each respective PafA aggregation temperature (Tagg) as EC_{50} values.

2.3 In-Silico characterization of PafA binders

A major obstacle in *Mtb* structural biology is the frequent insolubility of its recombinant proteins, a limitation affecting approximately two-thirds of the recombinant proteome. This challenge has biased the available structural data toward homologs from more tractable model organisms or the use of engineered mutants designed for enhanced solubility. The *Mtb* PPS proteome also faces the same limitations. In this study, we similarly identified and compiled available resolved structures of actinobacterial homologs of the *Mtb* PPS proteins. The suitability of these homologs as structural proxies was rigorously assessed through sequence and structural alignment analysis with their native *Mtb* PPS equivalents, providing a validated foundation for subsequent computational ligandability assessments.

2.3.1 Sequencing analysis of *Mtb* PPS proteins against resolved PPS homolog models

Sequential identity between the *Mtb* H37Rv PPS proteins and their resolved actinobacterial homologs was determined using the NCBI BLASTp database. Specifically, the sequences of *Corynebacterium glutamicum* PafA (cgPafA) (PDB: 4BJR) and *Acidothermus cellulolyticus* Dop (acDop) (PDB: 7OXY) were aligned against their respective *Mtb* counterparts, PafA (mtPafA; GenBank: CCP44872.1) and Dop (mtDop; GenBank: CCP44887.1). Both structural and docking analyses were performed on the monomeric unit of cgPafA, a form that is considered the catalytically active species, in contrast to the dimeric form in which the structure was resolved. Subsequently, structural alignments of these homologs were performed against their corresponding, high-confidence AlphaFold2-predicted models of the *Mtb* proteins (PafA: AF-C6DPW7-F1; Dop: AF-P9WNU8-F1) (Jumper *et al.*, 2021; Fleming *et al.*, 2025). All structural alignments and subsequent molecular modelling analyses were conducted using the Schrödinger Maestro suite (v2022-1), under an academic license provided by the Centre for High Performance Computing (Cape Town, South Africa).

2.3.2 In vitro PafA ligands docking analysis

Following model validation, cgPafA (PDB: 4BJR) was employed for molecular docking (Barandun *et al*., 2013). The screening focused on eight small-molecule ligands from the ChemBridge DIVERSet chemical library, which had been pre-screened to exclude compounds with aggregation-promiscuous tendencies. Given the structural similarity between the catalytic sites of PafA and Dop (Özcelik *et al*., 2012), acDop (PDB: 7OXY) was also included in the screening for comparative analysis (Cui *et al.*, 2021). All protein structures were prepared for docking analysis using the Protein Preparation Wizard within the Maestro suite with default parameters. Correspondingly, ligand preparations were performed using the LigPrep module. This included the set of eight small-molecule compounds (A2, A3, B7, C9, C10, C11, D2, D3) as well as the bona fide ligands ATP and ADP, and the short peptide substrates GGE and KGGE. Default settings were also applied for all ligand preparation steps.

Receptor grids for molecular docking were generated for the prepared protein structures. For the native cgPafA and acDop, grids were defined at three key sites: the ATP-binding site, the C-terminal Pup glutamine/glutamate binding site ($Pup^{E/Q}$ site), and the N-terminal Pup binding sites (comprising the IQK motif and alpha-helices 1 and 2). To investigate the impact of the *Mtb* Pup sequence, a hybrid model was generated—designated mtPafA—by mutating the Pup-binding groove of cgPafA to match that of *Mtb* (the pup binding groove of PafA being variable between *Cgl* and *Mtb*). Four corresponding receptor grids were generated for this mutant: $mtPup^{E}$, $mtPup^{IQK}$, mtPupα1, and mtPupα2. All ligands were docked into each of these receptor grids using the Glide module in Extra Precision (XP) mode. The resulting poses were subsequently subjected to binding affinity refinement via Prime MM-GBSA calculations. Ligands were ranked based on a consensus of their docking score, Glide Emodel, and the MM-GBSA-derived ΔG_{bind} value. To filter out weak binders, a docking score threshold of -4.0 kcal/mol was applied.

Molecular dynamics (MD) simulations were conducted for 100 ns on multiple states of cgPafA and mtPafA to characterize their dynamic behaviour. The simulated systems included the apo enzyme, complexes with catalytic components (cgPafA+Mg^{2+}, cgPafA-ATP+Mg^{2+}), and complexes with various ligands (A3, D2, C10, GGE), both in the presence and absence of Mg^{2+}. Additionally, the mutant mtPafA was simulated in its ATP+Mg^{2+} state and in complex with ligand C9. All simulations were performed with 20 mM $MgCl_2$ to mimic the *in vitro* assay

conditions, as previously described (Chetty *et al.*, 2025). Post-simulation binding free energies were calculated for each system using the MM-GBSA method. The analysis was performed on 901 frames per trajectory (frames 100 to 1000, inclusive), and the mean values for the total binding free energy (ΔGbind) and its key components—including Coulombic (ΔGcoulomb), covalent (ΔGcovalent), hydrogen bonding (ΔGHbond), lipophilic (ΔGlipo), packing (ΔGpacking), generalized Born solvation (ΔGsolvGB), and van der Waals (ΔGVdW) contributions—were computed and reported.

2.3. Effect of *In vitro* ligands in the mycobacterial PPS

A bacterial proliferation assay was conducted to evaluate the impact of the PafA candidate ligands on the growth of mycobacteria. To assess the potential effect of inhibiting the PPS system, growth was monitored under both standard and oxidative stress culture conditions.

2.3.1. *In vitro* ligands effect under standard growth conditions

Initially, a single colony of *Msm* mc^2 155 cells was inoculated into 10 mL Middlebrook 7H9 broth medium (supplemented with 0.2% Tween 80, 0.05% glycerol, 0.085% NaCl) and grown overnight at 37 °C, 150 rpm. The culture was then diluted 1:100 into the same fresh medium, assessed spectrophotometrically for initial OD_{600nm} (0.05 - 0.1), and the diluted culture subsequently transferred into sterile 96-well plates (Wuxi NEST Biotechnology Co., Ltd., Wuxi, China). Each compound was subsequently applied to the respective cell culture per well to 10 µM final concentration (100 µL final volume; n=3). Controls bithionol, ST1926, and DMSO control, were also included at their respective concentrations (n=3). The cell-compound solution plate was assessed at OD_{600nm} (0 h), and the plate was incubated at 37 °C for 24 h. Following incubation, the plate was spectrophotometrically analyzed for bacterial growth at each tested concentration, and the resulting growth data were fed into the GraphPad Prism 9.2.0.332 software for statistical analysis.

2.3.2. *In vitro* ligands effect under oxidative stress growth conditions

Oxidative stress was selected for analysis of the ligands on mycobacterial growth under stressful conditions. Oxidative stress upregulates the PPS to mount an effective stress response, thereby creating a physiological context in which PPS inhibition is likely to manifest as an anti-proliferative effect. The assay was performed as previously described, using Middlebrook

medium supplemented with additional sodium nitrite ($NaNO_2$). To establish the appropriate oxidative stress level, *Msm* mc² 155 cells were first treated with a dilution series of $NaNO_2$ (80–0.625 mM) alongside a kanamycin control (10 µg/mL). Based on the resulting susceptibility profile, a sub-lethal $NaNO_2$ concentration was selected. Cells were then cultured at this concentration in the presence of the candidate PafA ligands to evaluate their specific effect on proliferation under nitrosative stress.

CHAPTER 3. RESULTS

The integral ligase PafA, a key component of the mycobacterial pupylation cascade, was investigated as a potential target for therapeutic intervention. The protein was recombinantly expressed and purified to homogeneity. A screen of a small-molecule library against purified PafA was conducted to identify binding ligands. Positive hits from this screen were further characterized using biophysical and biochemical assays to determine their effects on PafA stability and enzymatic activity.

3.1 Recombinant expression and purification of pupylation proteins

3.1.1 GroEL-GroES-assisted expression of recombinant PafA

To facilitate soluble recombinant expression, PafA was co-expressed with the GroEL/GroES chaperone system in *E. coli*. BL21(DE3) cells were first transformed with the pGro7 vector (encoding the chaperones), followed by transformation with the pET24b(+)_PafA vector. Following sequential transformation, the co-expression of soluble PafA with the chaperones was confirmed by immunoblot analysis (Figure 3.1).

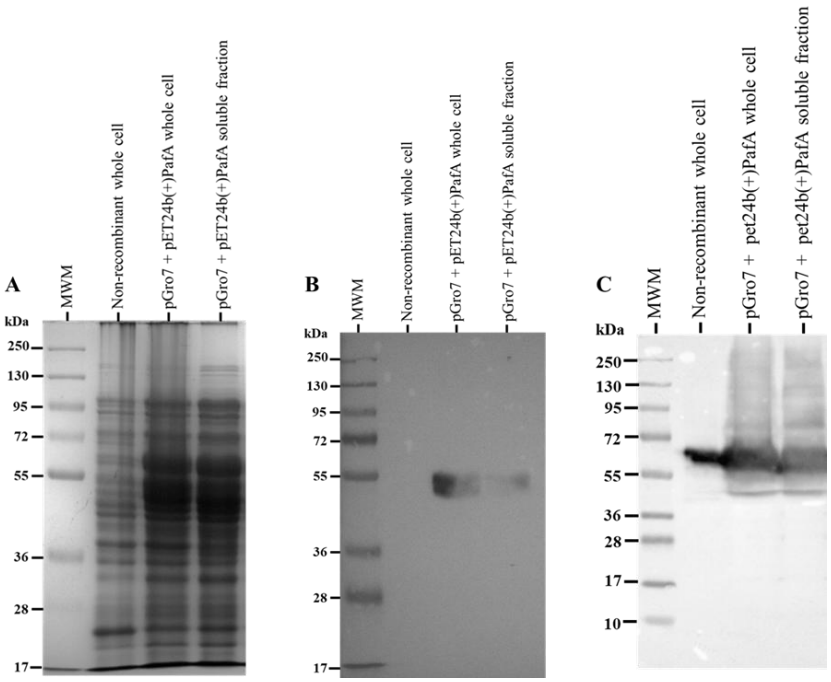

Figure 3.1 Soluble co-expression of recombinant PafA with the GroEL in *E. coli.* Treated samples were separated by (**A**) 10% SDS-PAGE gel and transferred onto 0.45 µM nitrocellulose by semi-dry transfer (10 V, 1 hour). Immunoblots were probed with **(B)** mouse monoclonal anti-6xHis antibodiesand goat monoclonal anti-mouse secondary antibodies and (**C**) mouse monoclonal anti-GroEL antibodies (cat no.:AD1-SPS-870) and Goat monoclonal anti-mouse secondary antibodies. Detection was facilitated by Clarity enhanced chemiluminescence (ECL) reagent.

As shown in Figure 3.1, the non-recombinant BL21(DE3) control strain grown with 0.5 mg/mL L-arabinose showed no detectable expression of 6xHis-tagged or GroEL recombinant proteins (Lane 2). In contrast, the dual recombinant strain, transformed with both pGro7 and pET24b(+)_PafA, exhibited a distinct change in the protein profile when grown under the same conditions. Two prominent protein bands were observed at approximately 54 kDa and 58 kDa. Anti-6xHis immunoblotting identified the ~54 kDa band as the recombinantly overexpressed 6xHis-PafA (Figure 3.1B, Lanes 3-4). The apparent molecular weight of this band was consistent with the 54 kDa size determined for the homologous resolved *Cgl* PafA monomeric structure (PDB: 4B0T; Özcelik *et al*., 2012).

Probing of the membrane with an anti-GroEL antibody confirmed the concomitantly overexpressed 58 kDa protein as GroEL, demonstrating L-arabinose induction of the pGro7 chaperone system (Figure 3.1C, Lanes 3-4). The relative increase detected in the recombinant strain (Lanes 3-4) contrasted with the endogenously detected GroEL in the non-recombinant control (Lane 2), confirming the overexpressed protein was vector-derived.

3.1.2 Recombinant PafA purification

Soluble *Mtb* PafA purification was achieved via Ni^{2+}-NTA IMAC (Figure 3.2). Incubation of the soluble protein fraction (Lane 2, Figure 3.2) with the resin resulted in the complete depletion of PafA from the flow-through (Lane 3), confirming efficient binding to the column. Elution of bound protein yielded moderately pure recombinant PafA (Figure 3.2A, Lanes 7-9), which co-purified with a single 73 kDa contaminant that was not detected during PafA immunoblot analysis (Figure 3.2B, Lanes 7-9). Examination of the purified protein (from Lane 7) under non-reducing conditions by SDS-PAGE (Lane 11) detected two closely migrating bands near 55 kDa (see Appendix Figure S1 Lane 12), both of which were confirmed to be PafA by immunoblotting. Notably, a third, faint high-molecular-weight band (>250 kDa) was also detected, suggesting the presence of a PafA oligomer that was stable under non-reducing treatment conditions

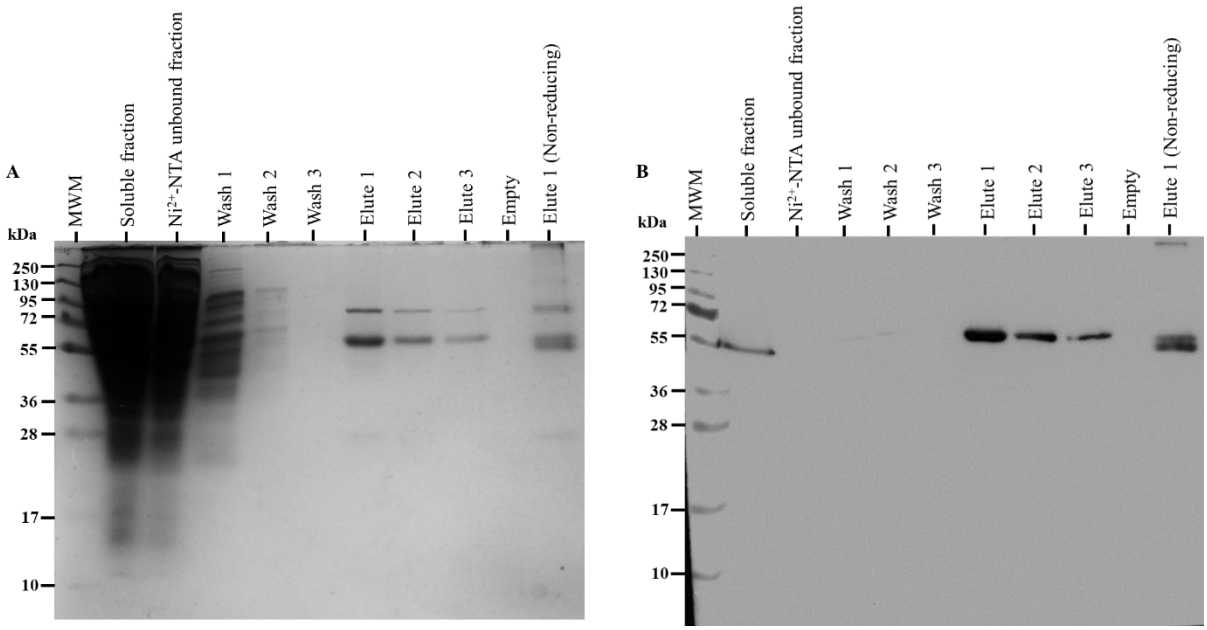

Figure 3.2. **Small-scale recombinant PafA purification A**. PafA Immobilised metal affinity chromatography (IMAC) treated samples were separated by (A) 12.5% SDS-PAGE gel, transferred onto 0.45 µM nitrocellulose membrane *via* semi-dry transfer **(**10 V**,** 1 hour**)**, and immunoblot detection was performed using (B) mouse monoclonal anti-6xHis antibodies and goat monoclonal anti-mouse secondary antibodies with clarity ECL substrate.

Following the optimization of recombinant PafA Ni^{2+}-NTA IMAC purification, the process was scaled up to 1-2 L expression cultures. The efficiency of this large-scale purification was assessed as shown in Figure 3.3. Purification of PafA reproduced the results observed at a small scale, with PafA co-eluting with a prominent contaminant of ~73 kDa that was not detected during PafA immunoblot analysis (Figure 3.3, Lanes 8-10). To investigate the properties of the purified PafA, purified PafA (sample from Lane 8) was analysed under different SDS-PAGE sample treatment conditions. Under standard reducing conditions, the protein's migration profile was identical regardless of whether the sample was boiled (compare Lane 11 to Lane 8, Figure 3.3A). Analysis under non-reducing conditions (without boiling and reducing agent) confirmed the presence of two distinct PafA bands at ~52 and ~54 kDa (Figure 3.3, Lane 13), consistent with previous observations. The inclusion of a boiling step in the non-reducing sample (Lane 14) did not alter this migration profile, indicating that heat denaturation alone does not affect the appearance of the two PafA bands that appear as a single band under standard reducing conditions (Lane 8).

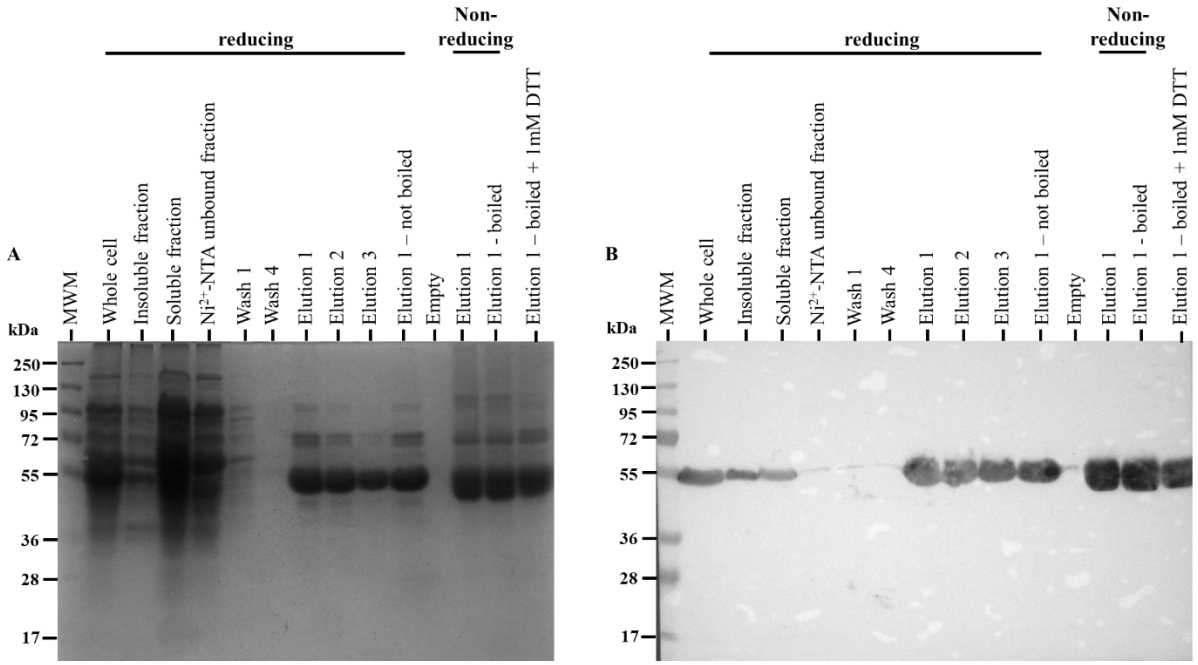

Figure 3.3. Scaled-up recombinant PafA purification. PafA Immobilised metal affinity chromatography (IMAC) treated samples were separated by (A) 10 % SDS-PAGE gel, transferred onto 0.2 μM nitrocellulose membrane via semi-dry transfer (10 V, 1 hour), and immunoblot detection was performed using (B) mouse monoclonal anti-6xHis antibodies and goat monoclonal anti-mouse secondary antibodies with clarity ECL substrate.

In contrast, treatment with the reducing agents (DTT or β-mercaptoethanol) caused a notable shift in the banding pattern (Lanes 8–10, 15). This shift occurred even when samples were not subjected to boiling prior to loading (Lane 11). The intensity of the lower 52 kDa band was substantially diminished, resulting in the merging of the two PafA bands towards a near-single predominant band at 54 kDa. This demonstrated that the redox environment, not thermal denaturation, was the primary determinant of PafA's double-banded electrophoretic mobility. Consequently, these experiments demonstrated that the redox state is a critical determinant of PafA stability. To maintain a consistent and defined structural state for all downstream applications, all subsequent purification and storage buffers were supplemented with the reducing agent DTT immediately following elution.

3.1.3 PafA biophysical characterization

To biophysically characterize and achieve high-purity PafA suitable for functional assays, the pooled, Ni^{2+}-NTA-purified fraction (stabilised with 1 mM DTT, pH 7.4) was further purified by SEC. The resulting chromatogram (Figure 3.4) revealed a prominent, broad elution peak

between 60 and 80 mL, consistent with the isolation of the PafA along with other protein contaminants in separate fractions.

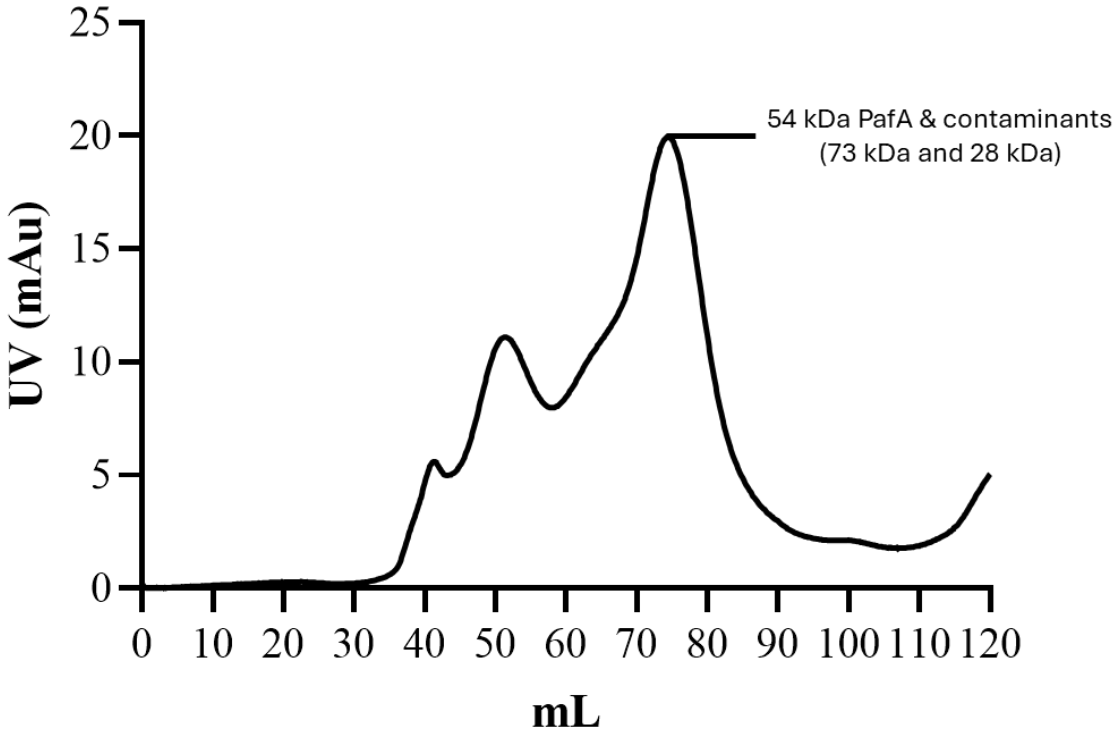

Figure 3.4. Recombinant PafA molecular exclusion chromatography (SEC) elution profile. Following purification by Ni^{2+}-NTA affinity chromatography, the recombinant PafA sample was clarified by centrifugation (20,000 × g, 10 min). The soluble supernatant was loaded onto a Sephacryl S-200 column pre-equilibrated with 1xEB (phosphate buffer, pH 7.4). Protein was eluted in the mobile phase as 2mL fractions at a flow rate of 0.5 mL/min. The chromatogram depicts the absorbance at 280 nm (mAu) versus elution volume.

The observed peak (60–80 mL) encompassed the expected elution volumes for the monomeric PafA (~68 mL) and the two major contaminants at ~73 kDa and ~28 kDa (approximately 65 and 74 mL), as confirmed by SDS-PAGE (Figure 3.5). Importantly, PafA and the contaminating species eluted near the BSA standard (65 ± 0.28 mL, Appendix Figure S2). The breadth of the peak, in contrast to the expected sharp resolution of SEC, indicated co-elution of multiple species. Furthermore, while all protein eluted within the resin's fractionation range (post-void volume: 37.47 mL), this range was too broad to resolve potential PafA dimers (ideally oligomers eluting in the void volume) from monomers.

Analysis of fractions corresponding to the broad peak (Fractions 30–40; 60–80 mL) confirmed monomeric PafA elution in this range, co-eluting with the anticipated 73 kDa and 28 kDa

proteins. A faint band, at the apparent expected size for PafA, was also observed eluting in earlier fractions (26–27), corresponding to the smaller peak at ~50 mL, which represented species of ≥300 kDa. The detection of faint PafA bands co-eluting with the ~73 kDa contaminant pointed to the presence of PafA in aggregated or complexed forms. Evidence for aggregation was provided by the significant quantity of PafA found in the insoluble pellet of the pre-analysed sample (Figure 3.5, Lane 2), indicating that aggregate PafA species may exist in solution.

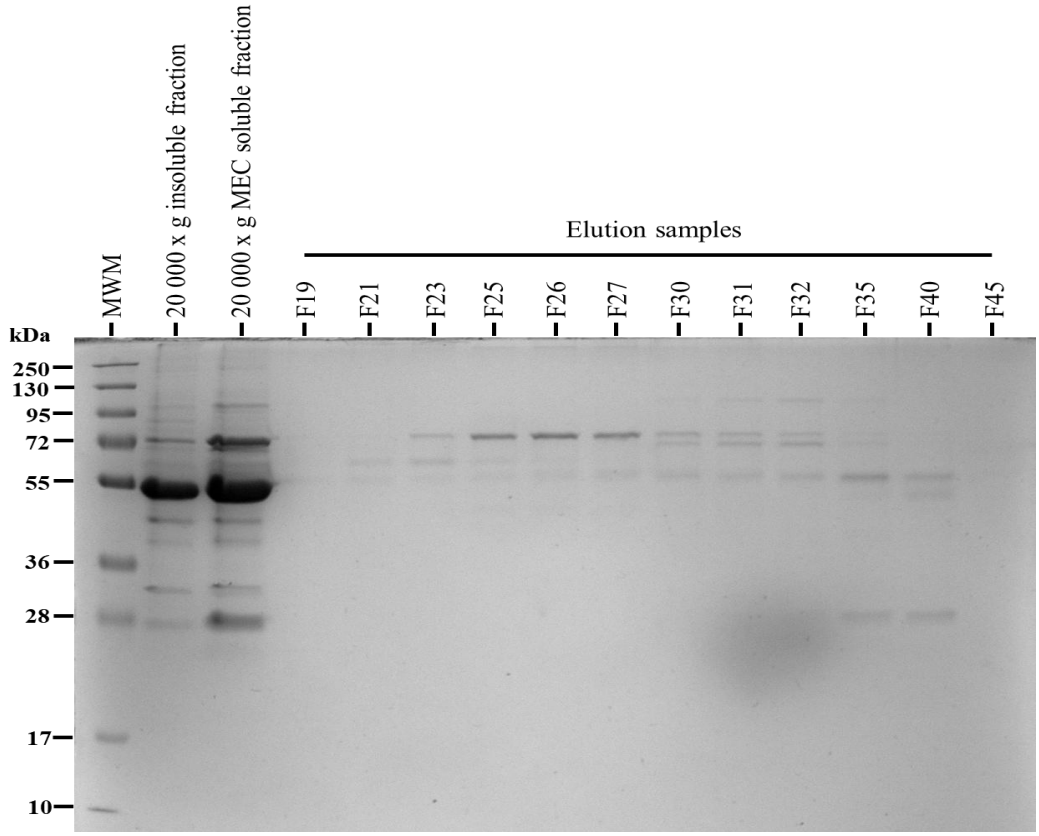

Figure 3.5. Analysis of recombinant PafA fractions from SEC. Ni^{2+}-NTA purification sample (Lane 3) was clarified by centrifugation (20,000 × g, 10 min) and loaded onto an equilibrated Sephacryl S-200 column. Select elution fractions (Lanes 4–15; 2 mL fractions) were analysed by 10% reducing SDS-PAGE, followed by Coomassie Brilliant Blue staining.

Overall, the eluted protein was dilute (Figure 3.5), requiring concentration via 10 kDa centrifugal filters. This step proved suboptimal, as PafA was incompatible with the filters, leading to substantial sample loss (see example in Appendix Figure S3). Consequently, while Sephacryl S-200 SEC identified that *Mtb* PafA predominantly exists as a monomer in solution, additional biophysical characterization methods were explored to characterize PafA while preserving protein concentration and integrity. Initial analysis via native PAGE assessed multiple Ni^{2+}-NTA-purified PafA stocks. The potential allosteric effect of ligand binding on

PafA oligomerization was probed using the native cofactor, Mg^{2+}. Analytical comparisons were performed with and without Mg^{2+}, which is typically used at high millimolar concentrations in functional assays, to determine its impact on potential quaternary structures (Figure 3.6). Following incubation with the cofactor Mg^{2+} (20 mM), both stored and freshly purified samples (lanes 2–5) revealed two prominent bands: a high-molecular-weight species (>250 kDa) and a faster-migrating band just above the 17 kDa marker. Importantly, immunoblotting confirmed that the high-molecular-weight band corresponded to PafA, indicating a homogeneous species in solution. In contrast, the lower band, which failed to transfer to the nitrocellulose membrane (evidenced by Ponceau S staining), was attributed to the ~28 kDa contaminant. The contaminant's relatively smaller, globular structure may have facilitated its deeper penetration into the gel matrix (assuming a pI < 8) during electrophoresis.

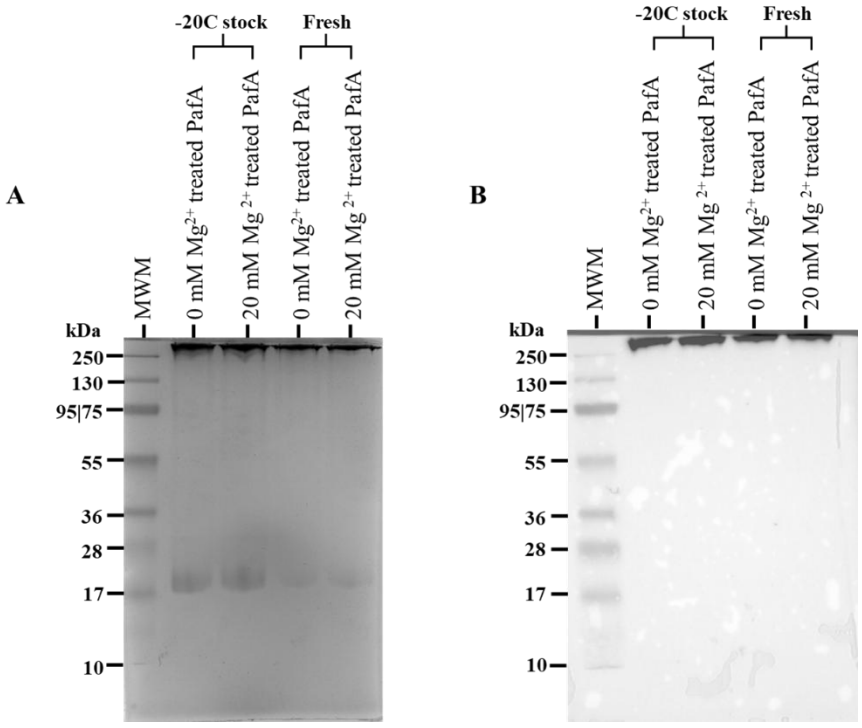

Figure 3.6. Native-state detection of recombinant *Mtb* PafA. Recombinant PafA was analyzed under non-denaturing conditions to assess its oligomeric state. (A) treated (20 mM Mg^{2+}, 1 h) and untreated Ni^{2+}-NTA purified PafA samples were resolved on a 10% non-denaturing (native) polyacrylamide gel. Proteins were then transferred onto a 0.2 µm nitrocellulose membrane via semi-dry blotting at a constant 10 V for 1 hour. (B) The native PafA species was detected by immunoblotting using a mouse monoclonal anti-6xHis primary antibody and goat anti-mouse HRP-conjugated secondary antibody, visualized with Clarity ECL substrate.

To further characterize PafA's oligomeric state, Ni^{2+}-NTA-purified protein was analyzed by SEC on a Sephadex G-75 column (fractionation range: 3–80 kDa). This resin provides higher resolution for distinguishing lower molecular-weight species compared to the S-200 resin used previously. SDS-PAGE analysis revealed that PafA did not elute in the void volume (4.5 mL),

indicating the absence of large oligomers or aggregates with a molecular weight ≥108 kDa under these conditions. Furthermore, the G-75 column successfully separated PafA from the primary 28 kDa and 73 kDa contaminants (the latter presumed DnaK), with both PafA and DnaK eluting in distinct, successive fractions (presumed DnaK in lanes 4–8; PafA in lanes 6–15, Figure 3.7).

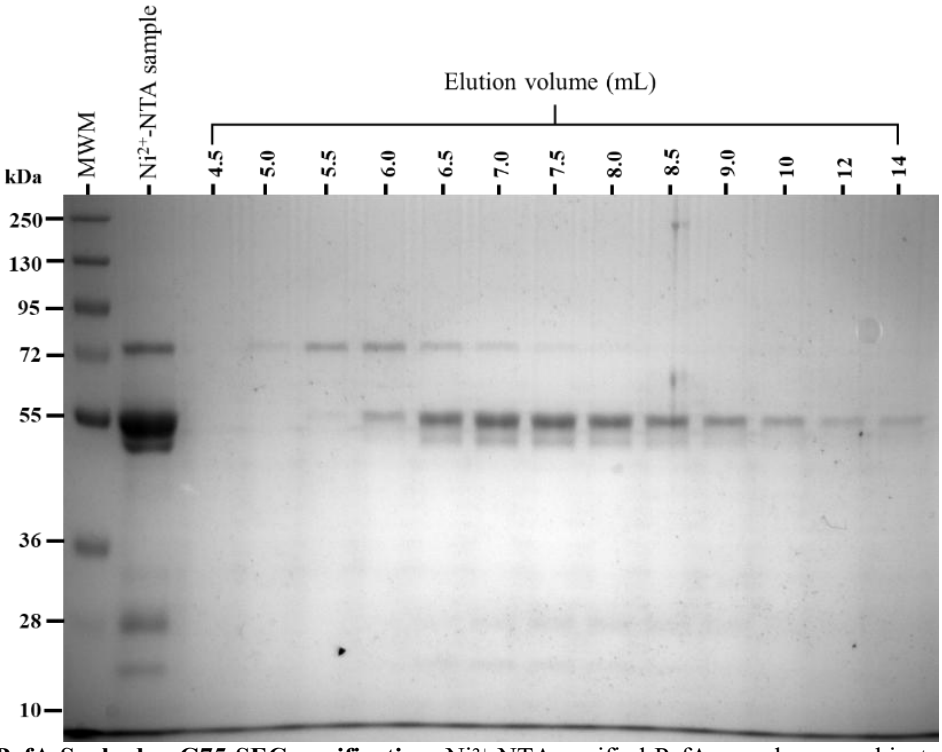

Figure 3.7. PafA Sephadex G75 SEC purification. Ni^{2+}-NTA purified PafA sample was subjected to a 1xEB (phosphate buffer, pH 7.4) equilibrated Sephadex G-75 column (15 mL bed volume). Selected elution fractions were resolved on a 10% reducing SDS-PAGE gel and subsequently visualized by Coomassie Brilliant Blue R-250 staining to assess purity and confirm the monomeric state of PafA.

Pure PafA was obtained in fractions 9-15. The elution volume of PafA was consistent with an exclusively 54 kDa monomer, a finding that contrasts with published reports for PafA homologs that exist heterogeneously as monomeric and dimeric forms in solution (Plourde *et al.*, 2024). Given the functional importance of Mg^{2+} for pupylation, it was hypothesized that the metal-free purification condition might influence oligomerization. To test this, PafA was incubated with 20 mM $MgCl_2$ and re-analysed by Sephadex G75 gel filtration (Figure 3.8). SEC of Mg^{2+}-treated PafA on the column revealed a complex elution profile. A significant portion of the protein was detected in the void volume (Figure 3.8, Lanes 3–5), indicating the presence of large complexes exceeding the column's fractionation range (≥108 kDa).

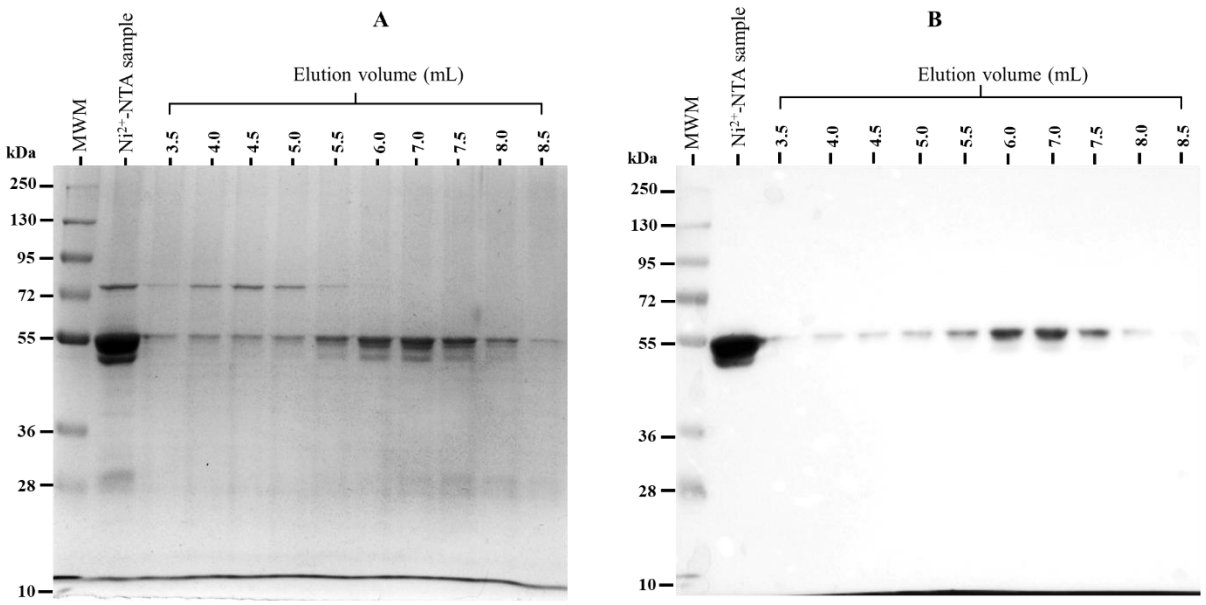

Figure 3.8. **Sephadex G75 SEC purification**. Ni^{2+}-NTA purified PafA sample (treated with 20 mM Mg^{2+} for 1 h) was subjected to a 1xEB (phosphate buffer, pH 7.4) equilibrated Sephadex G-75 column (15 mL bed volume). Selected elution fractions were resolved on a (**A**) 10% reducing SDS-PAGE gel and subsequently transferred onto 0.2 µM nitrocellulose membrane via semi-dry transfer (10 V, 1 hour), and (**B**) immunoblot detection performed using mouse monoclonal anti-6xHis 1° antibodies and Goat monoclonal anti-mouse 2° antibodies with clarity ECL substrate.

Analysis of this high-molecular-weight fraction revealed the co-elution of PafA with the 73 kDa DnaK protein, a finding that was absent in the apo-state PafA elution profile (Figure 3.7). The presence of both proteins in the void volume in approximately equimolar ratios suggested Mg^{2+} promotes the formation of a stable heterocomplex between the two proteins, a complex that would yield a molecular size of ~127 kDa, exceeding the resin's fractionation range. A second, more abundant PafA species eluted later, consistent with the elution of the monomeric species (Figure 3.8, Lanes 6–12). This monomeric fraction, confirmed by Plourde *et al*. (2025) as the catalytically relevant form for *Cgl* PafA, was subsequently isolated (Lane 8) for downstream ligand binding assays.

3.2 Identification and biophysical validation of novel PafA ligands

3.2.1 Identifying novel PafA ligands via the thermal shift assay

3.2.1.1 PafA TSA optimization

To assess protein stability and identify potential small-molecule ligands, a thermal shift assay was optimized for PafA (PafA TSA), as previously described in *section 2.2.1*. Initially, a thermal melt analysis of purified PafA (1 µg in 1xEB buffer) was conducted in the presence of 10X SYPRO orange dye from 25 to 99 ℃, yielding the thermal melt profile outlined in Figure 3.9.

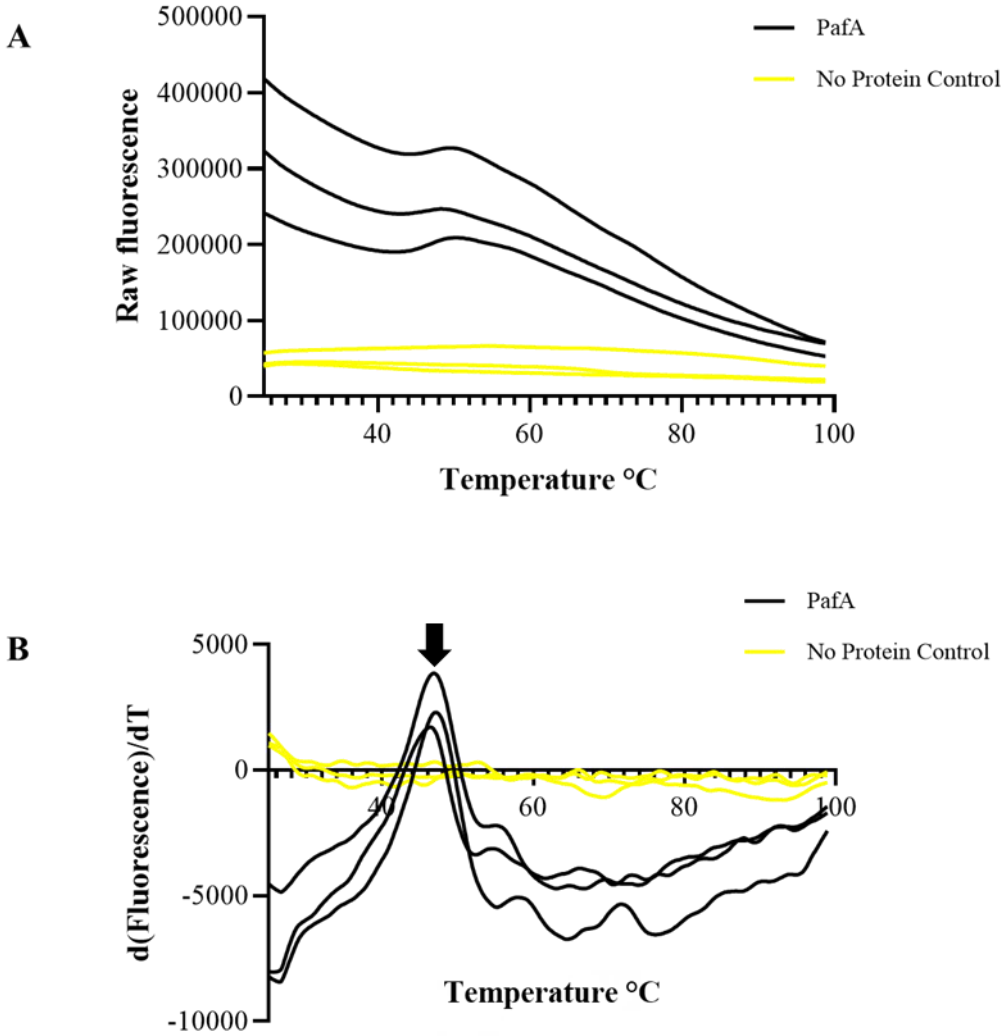

Figure 3.9. **Thermal melt analysis of purified PafA.** Protein thermal melt analysis was performed using 1 µg of purified PafA against 10X SYPRO orange dye at 0.05 ℃/s temperature increasing increments from 25 - 99 ℃. **(A)** The raw fluorescence curve for PafA (black) relative to the buffer-only control (yellow) and (**B**) the derivative melt curve, analysed via the Protein Thermal Shift Software to yield a melting temperature (Tm, shown by the arrow) of 46.80 ± 0.18 ℃ (mean ± standard error, n = 3).

Thermal melt analysis of purified PafA revealed intrinsic instability. The derivative melting temperature ($^{dFluorescence}/_{dTime}; Tm$) was calculated to be 46.80 ± 0.18 °C (Figure 3.9B), a value significantly lower than the bioinformatically predicted range of 55–60 °C. Furthermore, the raw fluorescence profile (Figure 3.9A) exhibited a high initial signal intensity, suggesting that the protein was in a partially unfolded state in solution. Specifically, the exposed hydrophobic residues bind SYPRO orange dye, hence the initial fluorescence. This was consistent with the observed behaviour of the protein, which consistently precipitated out of solution during storage. To address this, the buffer conditions were optimized by supplementing 1xEB with 1 mM DTT. The thermal melt profile of PafA in this stabilized buffer was subsequently re-analysed (Figure 3.10).

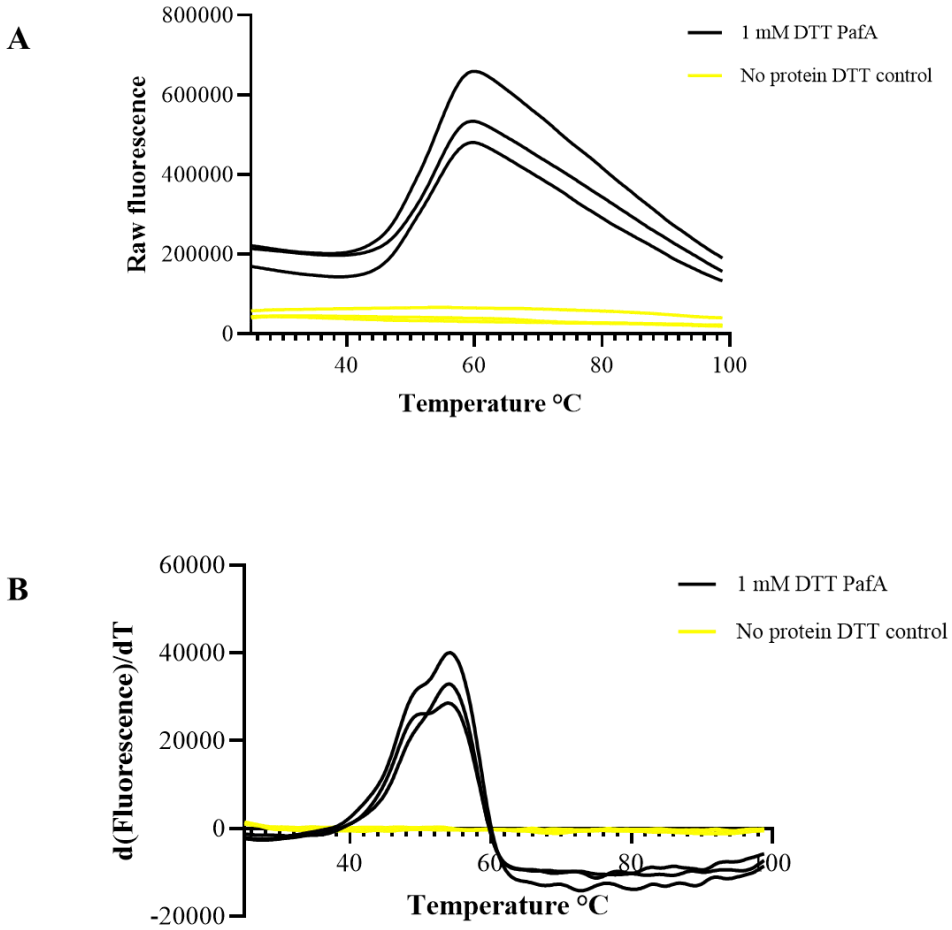

Figure 3.10. **Thermal melt analysis of stabilised PafA**. Protein thermal melt analysis was performed using 1 mM DTT-treated purified PafA (1 µg) against 10X SYPRO orange dye at 0.05 °C/s temperature increasing increments from 25 - 99 °C. **(A)** The raw fluorescence curve for PafA (black) relative to the buffer-only control (yellow) and (**B**) the derivative melt curve yielding a biphasic protein profile with calculated Tm's of 50.47 ± 0.51 °C and 54.06 ± 0.05 °C (mean ± standard error, n = 3).

The addition of 1 mM DTT to the storage buffer significantly altered the PafA thermal melt profile, indicating enhanced stability. The resulting curve exhibited a low initial fluorescence that remained stable from 25 to 40 °C, followed by a sharp, cooperative increase indicative of protein unfolding (Figure 3.10A). This classic melt profile signified that PafA was predominantly in a folded, stable state at the beginning of the assay. The calculated PafA T_m increased to 54.06 ± 0.05 °C (Figure 3.10B), representing a significant 7.26 °C increase from the previous T_m of 46.80 ± 0.18 °C (Figure 3.9). This substantial shift confirmed a significant increase in the protein's thermal stability (Figure 3.11), bringing the experimental T_m into closer agreement with the bioinformatically predicted T_m range of 55–60 °C.

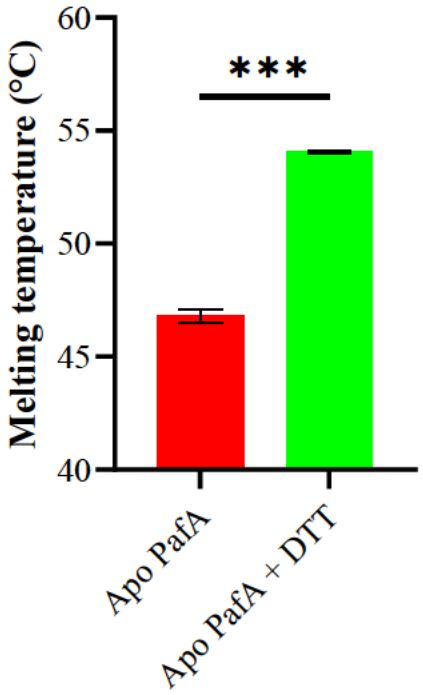

Figure 3.11. PafA stability buffer optimization. Analysis of purified DTT-treated PafA thermal melt reaction via the Protein Thermal Shift Software calculated PafA T_m at 54.06 ± 0.05 °C (mean ± standard error, n=3). Statistical significance was determined by an unpaired Welch's t-testing (*** $p < 0.001$).

The derivative melt curve for PafA revealed a biphasic unfolding transition, featuring a minor shoulder at approximately 50 °C and a dominant peak at 54.06 ± 0.05 °C (Figure 3.10B). While a double peak can be indicative of dimeric proteins and aligns with reports of PafA homologs existing in multiple oligomeric states (Plourde *et al.*, 2024), the data suggested a different cause. The two transitions were found to correspond to the co-purification of PafA in two distinct monomeric forms: an oxidized species and a reduced species, as reported earlier. The less stable oxidized population denatured first (T_m = 50.47 ± 0.51 °C), representing a 3.67 °C increase

from the previous T_m of 46.80 ± 0.18 °C (Figure 3.9), followed by the more stable reduced population (T_m = 54.06 ± 0.05 °C). To confirm that the reduced form was responsible for the higher-T_m peak, stored PafA was buffer-exchanged into fresh buffer containing 1 mM DTT and re-analysed. The resulting melt profile (Figure 3.12) depicted a single, consolidated peak at the higher melting temperature, confirming that the heterogeneity was due to redox state and not oligomerization. Analysis of the single-phase thermal melt profile yielded a melting temperature (T_m) of 54.70 ± 0.35 °C for PafA (Figure 3.12). This single, cooperative transition indicates that the equilibrium in solution was shifted overwhelmingly toward the stable, reduced monomeric state, effectively depleting the population of less stable oxidized species.

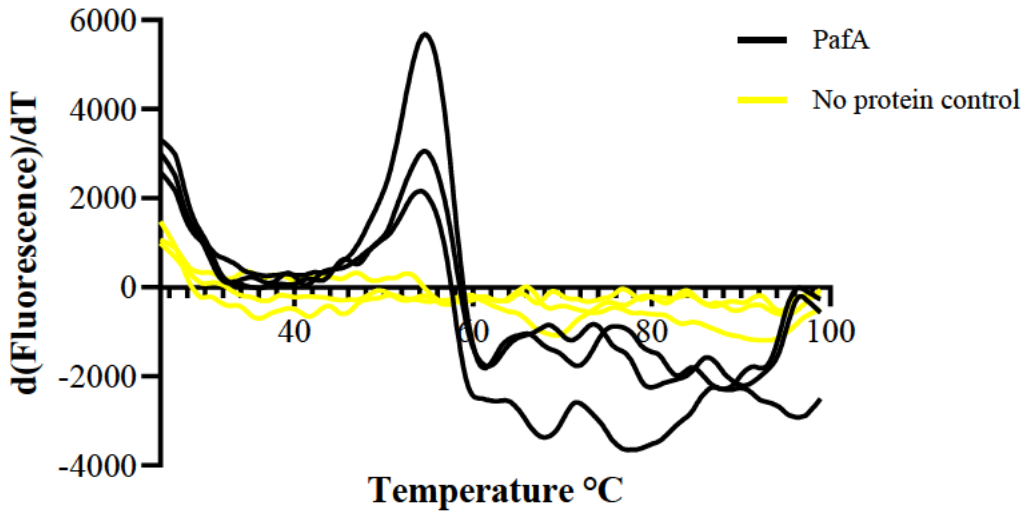

Figure 3.12. **PafA oligomeric state examination via the thermal melt assay.** Protein thermal melt analysis of 1 mM DTT-treated PafA was performed using 1 µg purified PafA against 10X SYPRO orange dye at 0.05 °C/s temperature increasing increments from 25 - 99 °C. Analysis of purified PafA thermal melt reaction via the Protein Thermal Shift Software resulted in the classical protein melt curve (**A**, Black) relative to the control (yellow) and the calculated derivative protein melting temperature (T_m) for PafA at 54.70 ± 0.35 °C (mean ± standard error, n = 3), illustrated by the peak observed on the derivative curve.

Despite this single, cooperative thermal melt transition under reducing conditions (1 mM DTT), PafA was found to occasionally populate two distinct monomeric states in solution, with the reduced 54 kDa monomeric state being the dominant form (Figure 3.7-3.8). These results indicated that the addition of the reductant shifts the conformational equilibrium predominantly toward the reduced monomeric state. However, the persistence of multiple monomeric forms under these conditions demonstrates that PafA exists as a dynamic monomeric ensemble in solution under these conditions. With this optimized and reproducible system, a PafA TSA was established. To validate the assay, the bona fide PafA ligand ATP was employed as a positive control to demonstrate the PafA T_m shift upon binding as a function of increased protein

stability. To do this, the optimised PafA thermal melt reaction was adopted to optimize and validate a PafA TSA, using ATP in the presence and absence native cofactors (Figure 3.13).

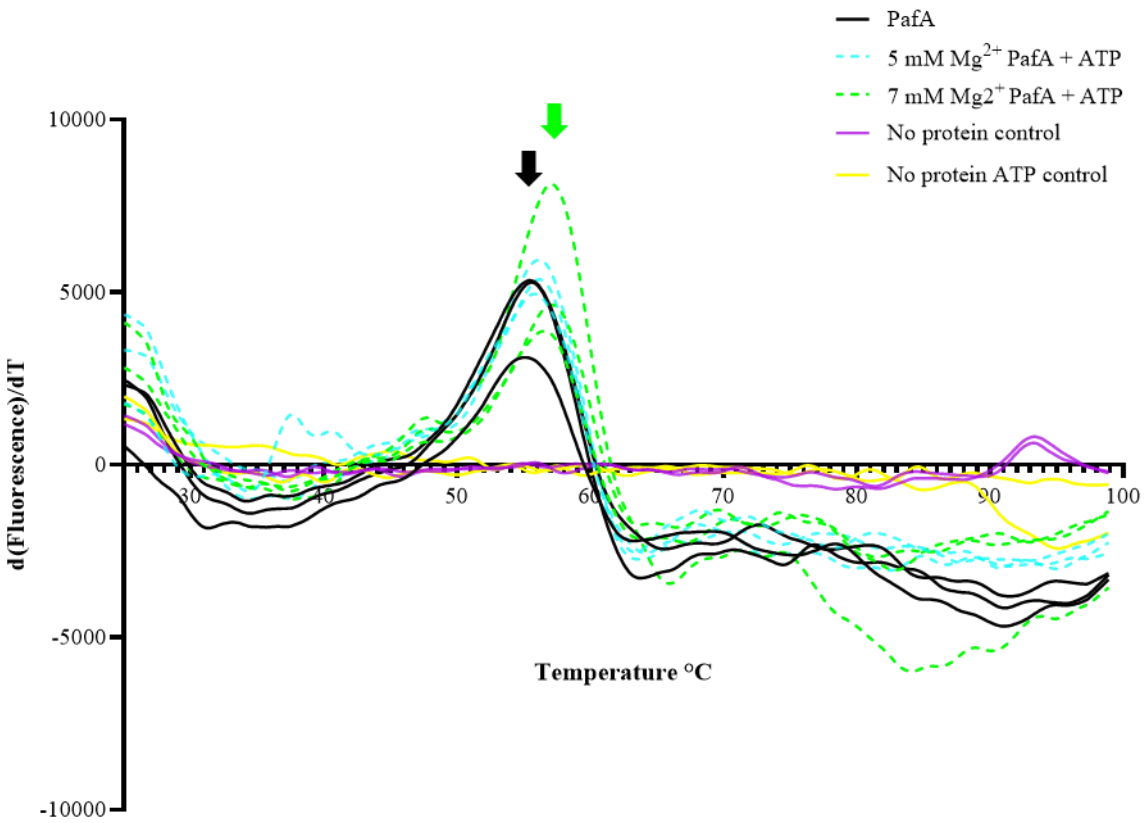

Figure 3.13. **Effect of Mg^{2+} and ATP on PafA thermal stability.** The thermal stability of PafA was assessed using the optimized thermal melt assay. PafA (1µg) was pre-incubated with Mg^{2+} for 1 h, followed by a 1 h incubation with 2.5 mM ATP at room temperature (RT). PafA thermal denaturation was subsequently monitored from 25 - 99°C at a rate of 0.05°C/s in the presence of 10X SYPRO Orange dye. Melt curves were analysed to determine the derivative PafA T_m (shown by arrows) for each respective treatment(n=3)

Specifically, the crystallized PafA homolog (PDB: 4BJR; Barandun *et al*., 2013) reveals the presence of magnesium (Mg^{2+}) and its critical role as a cofactor that coordinates the β- and γ-phosphates of ATP and key residues within the catalytic site, forming a prerequisite Mg-ATP complex for productive binding. To investigate whether Mg^{2+} is required to stabilize ATP binding, PafA was pre-incubated with Mg^{2+} to form the holo-enzyme prior to its use in the thermal melt assay with ATP. The results of this experiment are presented in Figure 3.13-3.14 and Table 3.1. Initial testing with 2.5 mM ATP in the optimized thermal shift assay yielded only a minor T_m increase of 0.26 °C, from 54.70 ± 0.35 °C to 54.97 ± 0.05 °C. This modest shift, while suggestive of binding, indicated that the assay required further refinement to robustly detect the ATP-PafA interaction.

Table 3.1. Optimization of the PafA thermal shift assay using native ligands

Compound	T_m (°C)	T_m shift	P-value	Significance (*P < 0.05*)
Apo PafA	54.70 ± 0.35	-	-	-
Apo PafA+ ATP	54.97 ± 0.05	0.26	0.9390	No
Holo^{5mM} PafA	55.40 ± 0.18	0.70	0.2354	No
Holo^{5mM} PafA + ATP	56.04 ± 0.13	1.33	0.0066	Yes
Holo^{7mM} PafA	55.45 ± 0.22	0.74	0.1855	No
Holo^{7mM} PafA + ATP	56.93 ± 0.20	2.22	<0.0001	Yes

Statistical analysis via One-way ANOVA multiple comparisons testing (T_m as mean T_m ± standard error; n = 3)

PafA thermal shift analysis revealed that PafA stability is specifically enhanced by the Mg-ATP complex. No significant change in the PafA T_m was detected upon the independent addition of Mg^{2+} (5 or 7 mM) or ATP (2.5 mM). However, a statistically significant increase in PafA T_m was observed when ATP was supplemented with either 5 or 7 mM Mg^{2+}. Specifically, PafA T_m significantly shifted from 54.70 ± 0.35 °C for the apo protein to 56.04 ± 0.13 °C with 5 mM Mg^{2+} (ΔT_m = +1.33 °C) and to 56.93 ± 0.20 °C with 7 mM Mg^{2+} (ΔT_m = +2.22 °C) (Figure 3.14). This demonstrated a Mg^{2+} concentration-dependent stabilization of PafA by ATP, revealing neither Mg^{2+} nor ATP can independently confer a significant stabilizing effect on PafA upon binding, but rather function synergistically to significantly increase PafA stability. As anticipated, ATP binding produced a characteristically positive shift in the PafA T_m (Figure 3.14), demonstrating the increased protein stability that occurs upon ligand binding and confirming the assay's functionality and veracity for subsequent PafA ligand screening.

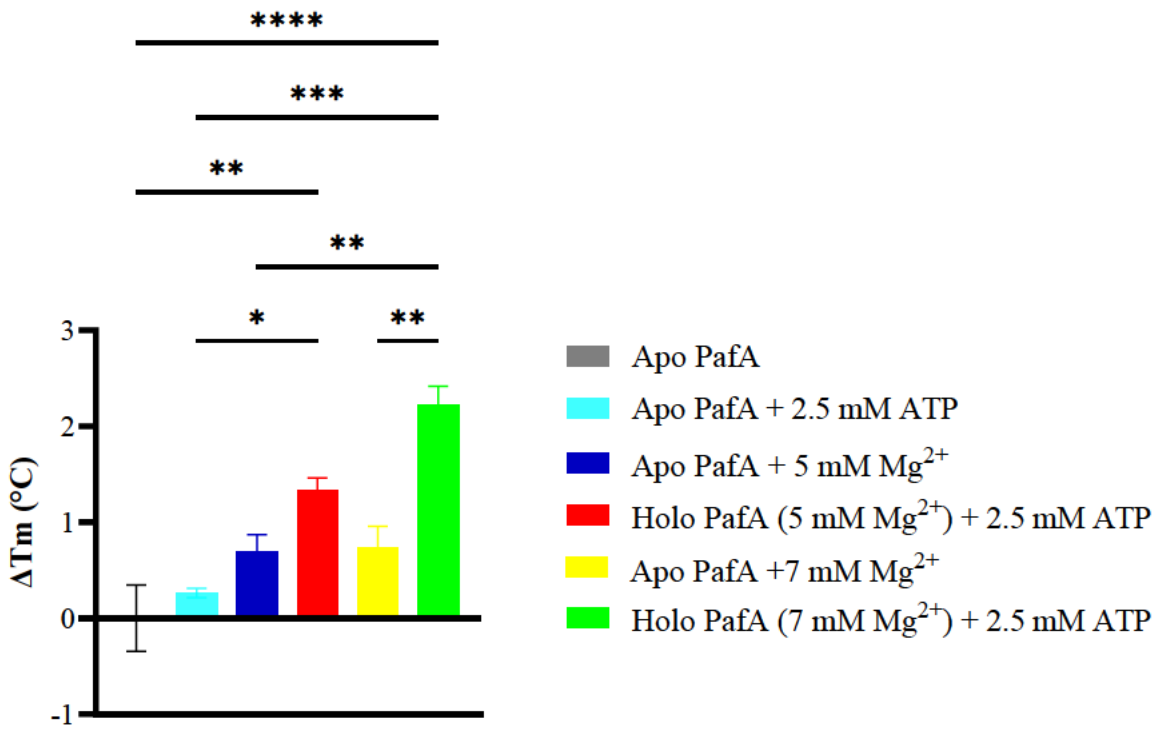

Figure 3.14. Thermal stabilization of PafA by ATP and Mg^{2+}. The optimized thermal melt assay was used to assess changes in PafA stability upon binding 2.5 mM ATP in the presence of Mg^{2+}. PafA ΔT_m (Tm ± standard error), was calculated for each treatment relative the control protein T_m 54.70 ± 0.35. Statistical significance was determined by one-way ANOVA with multiple comparisons (n=3; ****P < 0.0001, ***P < 0.001, **P < 0.01, *P ≤ 0.05).

3.2.1.2 PafA TSA: Small-molecule ligand screening

Following optimization, the PafA TSA was deployed to screen seven drug-like compounds from the ChemBridge DIVERSet library. These compounds had been identified as potential hits in prior ligandability studies. To validate the assay's performance, two known binders of mutated *Mtb* PafA (mtPafA) and *Cgl* PafA (cgPafA), bithionol and ST1926 (Li *et al.*, 2022), were screened (Figure 3.15-3.16). Interestingly, treatment with 10 µM bithionol not only decreased PafA's thermal stability but also induced a unique biphasic unfolding transition in its melt curves, contrasting with the enhanced stability previously observed by the ATP-Mg^{2+} complex. This unfolding profile, which curiously resembled that of unstable protein prior to optimization (Figure 3.9), suggested that bithionol exposure shifts the equilibrium of the two PafA monomeric forms in solution. Specifically, treatment with bithionol altered the distribution of PafA species, increasing the relative abundance of the less stable, oxidized form, which formed distinct species in solution in relative proportions (Figure 3.15). This suggested that bithionol binding may shift the conformational equilibrium or directly promote oxidation.

Furthermore, the elevated initial fluorescence signal indicated that the ligand-bound protein exists in a partially unfolded state prior to thermal denaturation.

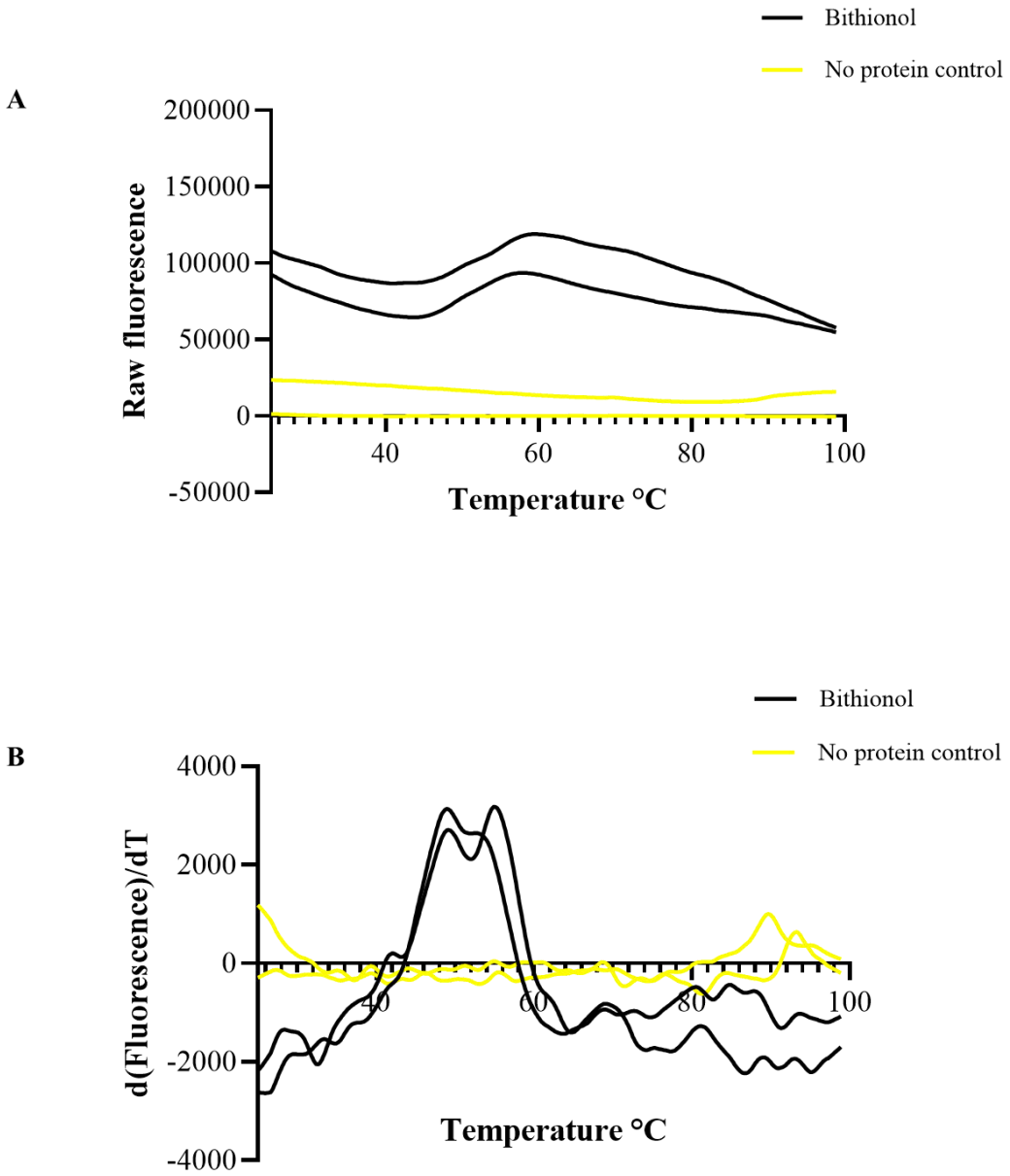

Figure 3.15 PafA thermal melt profile in the presence of bithionol. The thermal stability of PafA was assessed using the optimized PafA thermal shift assay. Mg^{2+} preincubated PafA (1µg) was incubated with 10 µM bithionol for 1 h at RT. PafA thermal denaturation was subsequently monitored from 25 - 99°C at a rate of 0.05°C/s in the presence of 10X SYPRO Orange dye. Melt curves were analysed to determine the derivative PafA T_m for each respective treatment (n=3).

The altered melt profile in the presence of bithionol indicated a ligand-induced effect, i.e., bithionol preferentially binds to the less stable PafA form, shifting the conformational equilibrium towards the less stable partially unfolded state. This resulted in a biphasic unfolding transition, characterised by the prominent presence of both the less stable PafA species (T_m = 49.48 ± 0.42 °C) and the more stable population with a T_m of 53.87 ± 0.68 °C, both of which

were not notably different from their respective unliganded biphasic control equivalents (T_{m1} = 50.47 ± 0.51 °C and T_{m2} = 54.06 ± 0.05 °C; Table 3.2).

This model was supported by the lower initial fluorescence intensity of the bithionol-treated protein sample compared to the unstable control (Figure 3.9A vs. 3.15A), suggesting that only the less stable subpopulation is partially unfolded and contributes to the elevated signal. To evaluate whether a ligand binding to a site other than the ATP site would similarly confer stability, the PafA TSA was subsequently applied to ST1926, a Pup^E groove-targeting compound (Li *et al.*, 2022), evaluating its impact on PafA conformational stability (Figure 3.16).

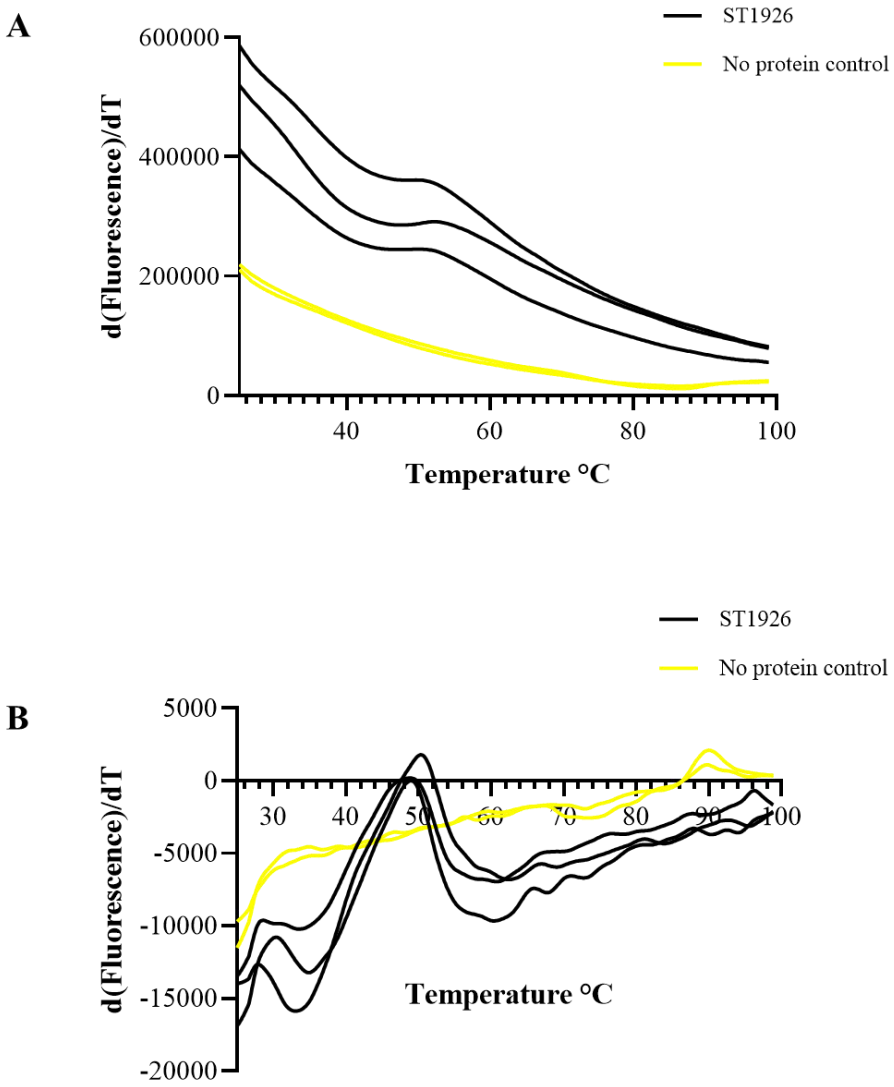

Figure 3.16. PafA thermal melt profile in the presence of ST1926. The thermal stability of PafA was assessed using the optimized PafA thermal shift assay. Mg^{2+} preincubated PafA (1µg) was incubated with 10 µM ST1926 for 1 h at RT. PafA thermal denaturation was subsequently monitored from 25 - 99°C at a rate of 0.05°C/s in the presence of 10X SYPRO Orange dye. Melt curves were analysed to determine the derivative PafA T_m for each respective treatment (n=3).

Similarly, ST1926 also significantly altered the proteins melt profile. In contrast to the stabilizing effect of the ATP-Mg^{2+} complex, ST1926 binding resulted in a pronounced reduction in the protein's T_m to 49.38 ± 0.47 °C relative to the established stable PafA T_m of 54.06 ± 0.05 °C. This induced decrease in PafA's T_m suggested the compound preferentially binds to and thermodynamically stabilizes a partially unfolded, less stable conformation. This binding event shifts the conformational equilibrium in solution, resulting in a population dominated by the destabilized, ligand-bound complex. This is supported by a high initial fluorescence observed in the PafA TSA (Figure 3.16A), suggesting that the PafA population in solution is dominated by species that adopt a partially unfolded state in solution, which is preferentially bound by the compound. Importantly, in this model, ST1926 binding does not cause a notable change in PafA T_m relative to the unliganded, less stable control equivalent (from 50.47 ± 0.51 °C to 49.38 ± 0.47 °C upon ST1926 binding). Unlike bithionol, ST1926 binding resulted in a single, cooperative unfolding transition (Figure 3.16B). This suggested the formation of a homogeneous complex, despite its reduced stability. The monophasic profile implied that ST1926 interacts with a specific, partially unfolded conformation of PafA, guiding its unfolding along a singular, well-defined pathway without accumulating detectable intermediate states.

Table 3.2. PafA thermal stability analysis in the presence of controls bithionol and ST1926

Compound	T_m (°C)	T_m shift	P-value	Significance (***P < 0.05***)
DMSO control	50.47 ± 0.51	-	-	-
ST1926	49.38. ± 0.47	- 1.09	0.3009	No
Bithionol (peak 1)	49.48 ± 0.42	- 0.99	0.2481	No

Statistical analysis via One-way ANOVA multiple comparisons testing (T_m as mean T_m ± standard error; n=3)

Despite no PafA thermal stabilization effects observed for ST1926 and bithionol, in fact exhibiting T_m shifts lower than the unliganded, less stable control protein, the change in the thermal melt profile, together with ~1 °C decrease in stability, provides strong qualitative evidence of ligand binding. This "destabilizing" effect indicates a binding mechanism distinct from that of the stabilizing Mg^{2+}-ATP complex. Specifically, both compounds appear to preferentially interact with a less stable, partially unfolded conformation of PafA, a preference most pronounced for ST1926. These results establish a valuable comparative framework; the destabilizing profiles of ST1926 and bithionol serve as functional counterpoints to the

stabilizing Mg^{2+}-ATP control, providing reference data for ligands that engage PafA through non-canonical, stability-reducing mechanisms.

Following validation with the control ligands bithionol, ST1926, and the Mg^{2+}-ATP complex, the optimised PafA TSA was employed to screen the seven small-molecule compounds from the ChemBridge DIVERSet collection for PafA binding. The resulting melt curves (Figure 3.17) were analysed for ligand-induced shifts in melting temperature (ΔT_m), reporting on changes to the protein's conformational stability. Analysis of the PafA thermal melt curves for the seven screened compounds revealed a consistent stabilizing effect (Figure 3.17). All seven compounds induced a positive shift in the protein's melting temperature relative to the DMSO control and 40 negative control compounds. Furthermore, whereas the control exhibited a biphasic unfolding profile, ligand-treated samples yielded predominantly near-monophasic transitions. Similar to ST1926, this shift to monophasic unfolding implied that the compounds bind PafA through a specific, well-defined interaction that commits the proteins in solution toward the more stable folded conformations, resulting in a uniform and cooperative denaturation process without populating significant intermediate states.

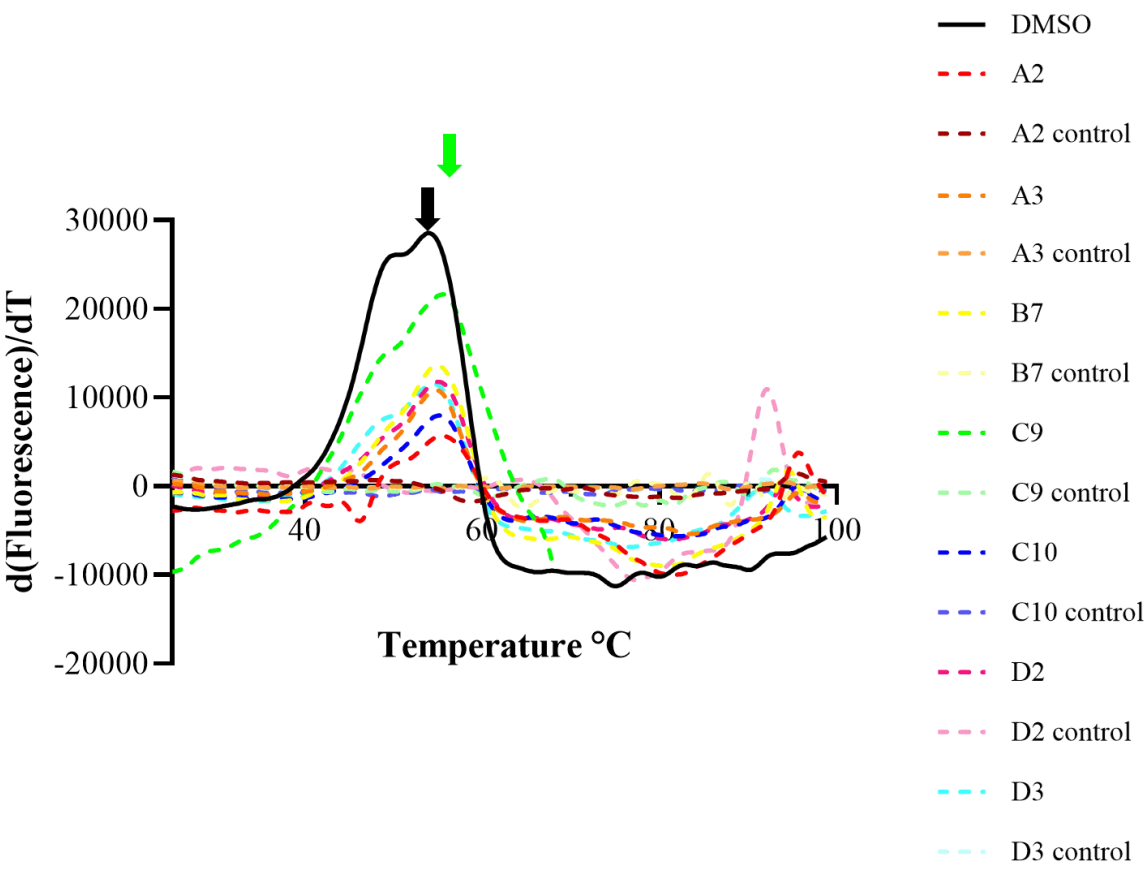

Figure 3.17. The identification of novel PafA ligands via the PafA TSA. The thermal stability of PafA was assessed using the optimized PafA thermal shift assay. Mg^{2+} preincubated PafA (1µg) was respectively incubated with 10 µM compounds for 1 h at RT. PafA thermal denaturation was subsequently monitored from 25 - 99°C at a rate of 0.05°C/s in the presence of 10X SYPRO Orange dye. Melt curves were analysed to determine the derivative PafA T_m for each respective treatment (n=3).

This simplification of the melt profile suggested that ligand binding selects for or induces a single, more stable PafA conformation in solution. The quantified T_m values for each compound, demonstrating the extent of this significant stabilization, are presented in Figure 3.18 and Table 3.3. The positive T_m shifts observed for all seven compounds demonstrated a consistent and significant stabilization of PafA, with ΔT_m values exceeding 0.7 °C for all compounds (Figure 3.18, Table 3.3). This outcome aligned with the stabilization conferred by ATP (Figure 3.13, Table 3.1) and is functionally distinct from the pronounced destabilization caused by bithionol and ST1926, revealing a divergent mode of action for the identified *in vitro* ligands.

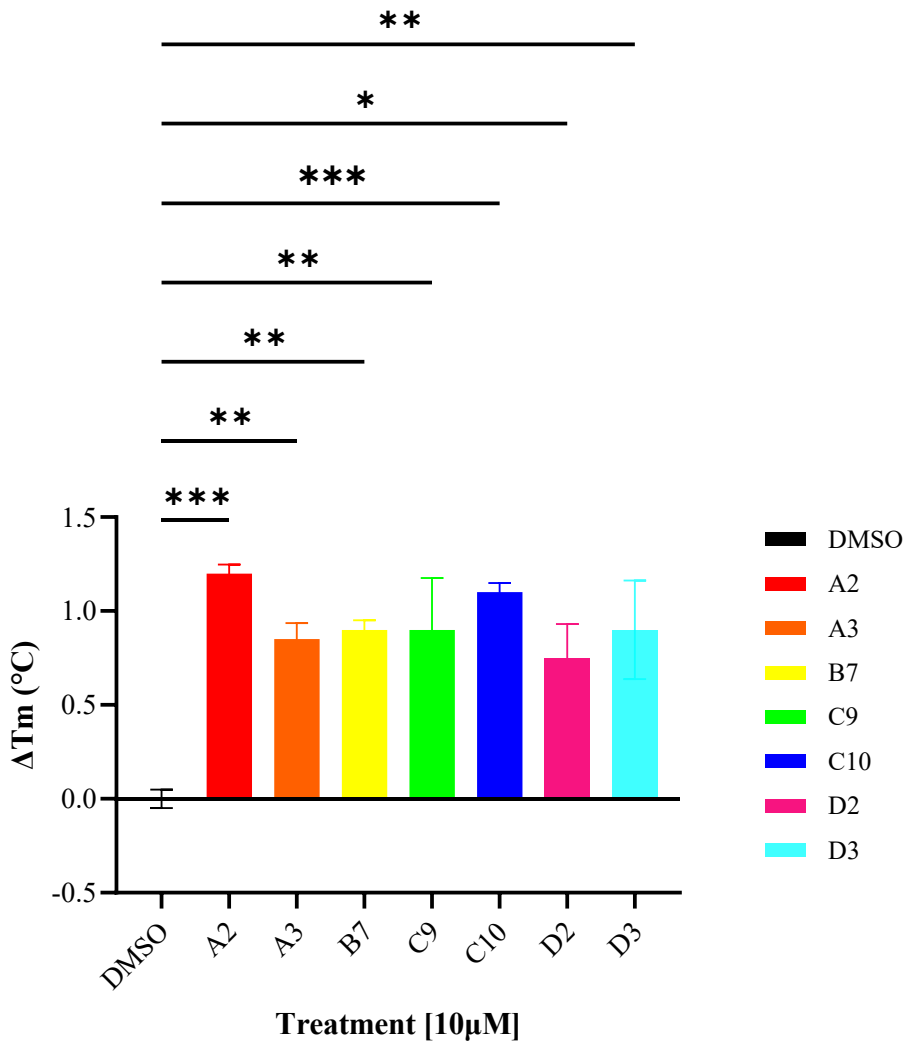

Figure 3.18. PafA thermal stability in the presence of small-molecule ligands. The optimized thermal melt assay was used to assess changes in PafA stability upon binding of the small-molecule ligands in the presence of 7 mM Mg^{2+}. PafA ΔT_m (T_m ± standard error), was calculated for each treatment relative the DMSO control T_m (54.06 ± 0.05 °C). Statistical significance was determined by one-way ANOVA with multiple comparisons testing (n=3; *** P<0.001, ** P<0.01, * P<0.05)

The screening identified two top-performing stabilizers: A2 and C10, which increased the PafA T_m by 1.20 ± 0.05 °C and 1.10 ± 0.05 °C, respectively (Table 3.3). Notably, their efficacy rivalled that of the natural ligand ATP (ΔT_m = 1.33 ± 0.13 °C with 5 mM Mg^{2+}). A further four compounds (A3, B7, C9, and D3) demonstrated strong stabilization, with ΔT_m values of ~0.9 °C (Table 3.3), indicating a consistent trend of enhanced conformational stability across the hit compounds. Following the establishment of *in vitro* binding for all seven compounds, the molecular mechanisms of action for each compound was subsequently sought. Computational modelling was employed to simulate interactions between PafA and the ligands, with the aim of elucidating their mode of engagement with the PPS via PafA.

Table 3.3. PafA thermal stability in the presence of selected small-molecule ligands

Compound	T_m (°C)	T_m shift	P-value	Significance (*P < 0.05*)
DMSO	54.06 ± 0.05	-	-	-
A2	**55.23 ± 0.05**	**1.20**	**0.0003**	**Yes**
A3	54.91 ± 0.09	0.85	0.0077	Yes
B7	54.96 ± 0.05	0.90	0.0049	Yes
C9	54.96 ± 0.28	0.90	0.0049	Yes
C10	**55.16 ± 0.05**	**1.10**	**0.0008**	**Yes**
D2	54.81 ± 0.18	0.75	0.0193	Yes
D3	54.96 ± 0.26	0.90	0.0049	Yes

Statistical analysis via One-way ANOVA multiple comparisons testing (T_m as mean T_m ± standard error; n=3)

3.2.2 *In vitro* ligands *in silico* validation

3.2.2.1 Sequencing and structural analysis of pupylation proteins

The sequential and structural conservation between *Mtb* PPS proteins and their resolved homologs was evaluated (Figure 3.19). Sequential alignment revealed moderate identity, with mtPafA sharing 54.2% identity with cgPafA and mtDop sharing 61.3% with acDop (see Appendix on secondary structure alignment; SSA alignment).

In contrast, structural analysis demonstrated a high degree of conservation in functional sites. The ATP-binding site (defined as residues within a 5 Å radius of the bound ligand) was 100% conserved between mtPafA and cgPafA. Conservation within the Pup-binding groove, however, was more variable. When analysed using the same spatial parameters, the C-terminal Pup^{E} binding site was 100% conserved, while the Pup^{IQK}, Pupα2, and Pupα1 sites showed conservation levels of 80% (16/20 residues), 57% (8/14 residues), and 47% (7/15 residues), respectively. This gradient of conservation across the Pup-binding groove was consistent with the inter-species variability previously reported by Li and coworkers (Li *et al.*, 2022).

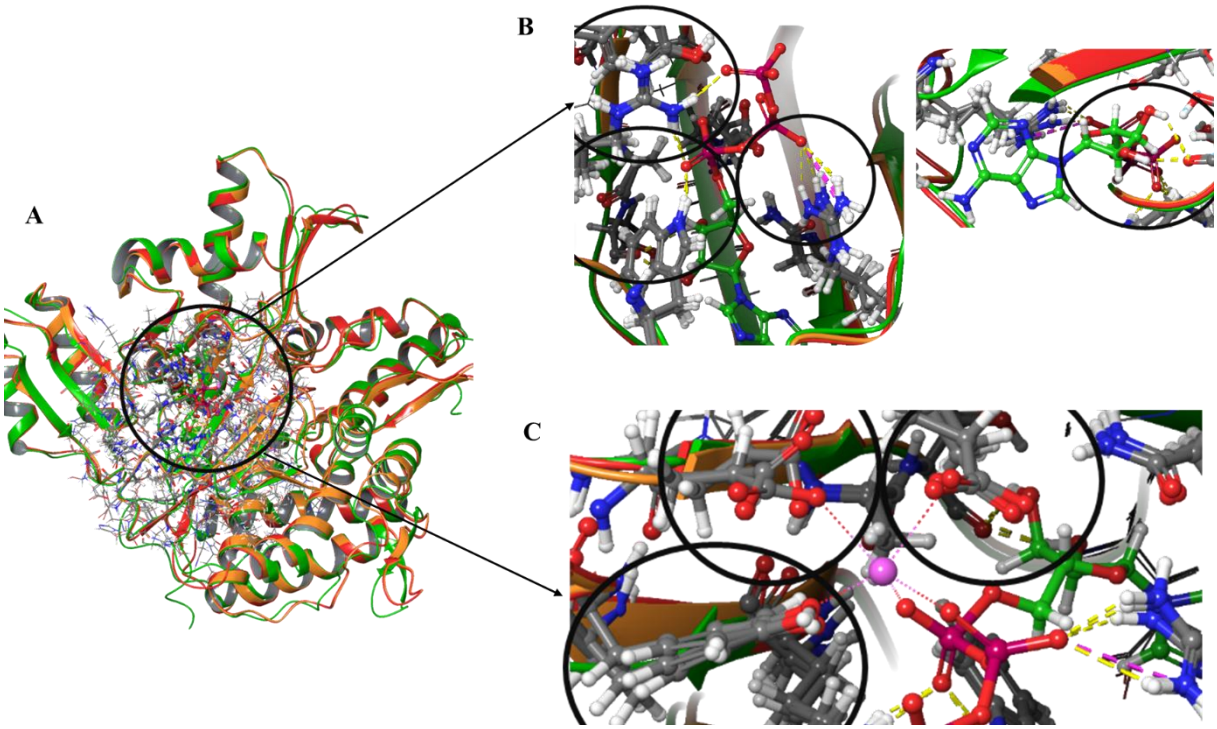

Figure 3.19. Structural conservation analysis of actinobacterial PafA homologs. The (A) PafA structures from *Corynebacterium glutamicum* (cgPafA, PDB: 4BJR; green), *Mycobacterium tuberculosis* (mtPafA, AF: C6DPW7-F1; red), and *Mycobacterium smegmatis* (msPafA, AF: A0QZ42-F1; orange) were aligned using Schrödinger Maestro (v2022-1). Receptor grids defining the ATP and Pup binding sites were subsequently generated based on the aligned structures. The (B) ATP-binding site was 100% conserved across homologs, with all key ATP-coordinating residues and their spatial orientation preserved. Similarly, the (C) Mg^{2+}-binding site showed high conservation, with two of its three coordinating residues conserved in the correct conformation. In contrast, the Pup-binding groove exhibited graded, region-specific conservation: 100% for the C-terminal PupE motif site, 80% for the Pup^{IQK} region, 57% for the Pup α-helix-2 binding domain, and 47% for the Pup α-helix-1 binding domain

3.2.2.2 Computational analysis of Mtb PafA in vitro ligands

To validate the target site with a non-native ligand, bithionol—a previously reported binder of the PafA ATP pocket (Li *et al*., 2022)—was docked into the cgPafA model. The docking scores were moderate, ranging up to -6.338 kcal/mol (Table 3.4), but successfully corroborated the binding site. Intriguingly, MM-GBSA analysis revealed that the calculated binding free energy (ΔGbind) was more favorable in the absence of Mg^{2+} (the same MM-GBSA metric yielding a record low of -6.03 kcal/mol). This suggested that the cofactor's absence conferred greater conformational flexibility to the ligand, allowing it to sample more optimal binding poses.

Using bithionol's docking profile as a benchmark, all eight candidate ligands were subsequently screened against the same site. Docking analysis of the eight ligands, including the natural substrates ATP and ADP, was performed for the ATP site of the structural homologs cgPafA

and acDop. Interestingly, all eight ligands demonstrated a higher binding affinity for the ATP site of cgPafA compared to the acDop ATP binding site (Table 3.4).

The presence of Mg^{2+} generally enhanced the binding affinity for both native substrates (ATP/ADP) and the eight tested *in vitro* ligands, as indicated by the overall lower (more negative) docking scores and free-binding energies (Table 3.4). However, this trend was not universal. Notably, ligands B7 and C9 exhibited more favourable docking scores in the absence of Mg^{2+} (-6.319 and -5.767 kcal/mol, respectively) than in its presence (-4.876 and -5.286 kcal/mol). Despite this, the more rigorous free-binding energy calculations Molecular Mechanics Generalized Born Surface Area (MM/GBSA) revealed that the binding of B7 and C9 was ultimately more thermodynamically favourable *with* Mg^{2+}, as reflected by their more negative MM/GBSA and Glide scores. This suggested that for these specific ligands, the presence of the co-factor stabilizes the complex in a manner not fully captured by the initial docking score.

Analysis of the eight ligands screened against the ATP site in the presence of Mg^{2+} identified three lead candidates—A3, C10, and D2—with docking scores and free-binding energies (ΔG_{bind}) comparable to those of the native substrates ATP and ADP, prompting their selection for further analysis. Notably, docking simulations of A3 in the absence of Mg^{2+}exhibited interesting results. Despite the A3 docking score being less favourable in the absence of Mg^{2+} (-6.7 vs -8.3 kcal/mol), the compound's final calculated ΔG_{bind} remained comparable (-24 vs -26 kcal/mol). This discrepancy between the initial pose prediction and the final energy estimation, coupled with the high potency of compound A3, prompted a focused analysis of the A3-PafA complex to elucidate the role of Mg^{2+} in PafA ligand binding. To explore alternative binding strategies by the ligands, the Pup-binding groove of cgPafA was screened. Structural analysis revealed this groove to be an extensive, solvent-accessible surface, in stark contrast to the deep, hydrophobic ATP-binding pocket (Figure 3.20). Given its topology, the Pup-binding groove was inferred to be able to accommodate ligands at multiple, non-overlapping sites.

Table 3.4. Ligand docking scores and MM/GBSA free-binding energies for *in vitro* PafA ligands docked into the ATP site with and without Mg^{2+} of PafA from *Corynebacterium glutamicum*

Compound	Mg^{2+} in the ATP site			No Mg^{2+} in the ATP site		
	Docking score (kcal/mol)	**Glide emodel (kcal/mol)**	**MM/GBSA ΔG_{bind} (kcal/mol)**	**Docking score (kcal/mol)**	**Glide emodel (kcal/mol)**	**MM/GBSA ΔG_{bind} (kcal/mol)**
ADP	-11.327	-123.739	-35.90	-8.942	-74.519	-39.02
ATP	**-11.278**	**-126.714**	**-47.51**	**-9.905**	**-90.005**	**-49.03**
Bithionol	-6.338	-41.55	-6.03	-	-	-
Bithionol	-	-	-	-5.877	-29.22	-62.16
Bithionol	-	-	-	-5.577	-26.34	-85.94
D2	**-8.592**	**-57.666**	**-42.3**	-4.388	-61.708	-35.68
D2	-8.533	-66.988	-45.59	-	-	-
C10	**-8.474**	**-63.234**	**-55.80**	-5.316	-41.316	-31.81
A3	**-8.338**	**-51.126**	**-24.23**	**-6.658**	**-49.788**	**-26.97**
A3	-8.243	-49.83	-25.01	-6.456	-44.269	-29.28
A3	-	-	-	-6.395	-51.499	-33.55
A3	-	-	-	-5.521	-47.059	-32.64
A3	-	-	-	-4.902	-37.65	-34.66
A3	-	-	-	-4.509	-37.798	-37.72
A3	-	-	-	-4.448	-42.02	-28.12
A3	-	-	-	-4.245	-34.717	-28.36
C10	-8.203	-58.771	-44.37	-	-	-
A2	-6.44	-5.909	-5.60	-4.709	-26.701	-9.03
A2	-6.34	-7.115	-12.23	-4.609	10.673	34.63
C11	-5.718	-42.574	-9.04	-	-	-
C11	-5.583	-43.448	-5.33	-	-	-
C9	-5.286	-35.547	-65.26	-5.767	-36.602	-35.21
C9	-4.969	-29.445	-54.48	-4.658	-42.23	-37.99
B7	-4.876	-13.673	-49.96	-6.139	-30.684	-34.73
B7	-4.808	-10.696	-33.46	-4.385	-22.684	-30.87
D3	-4.791	-47.513	-57.71	*-4.272*	*-51.5*	*-46.23*

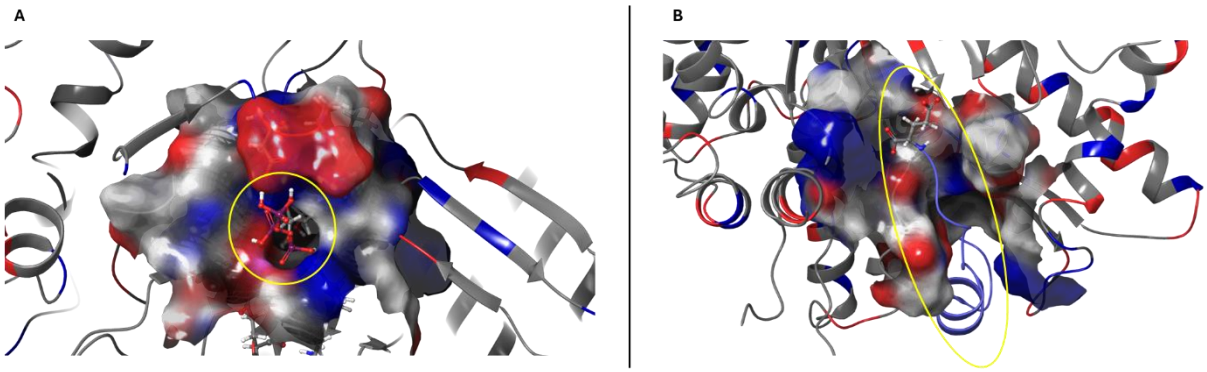

Figure 3.20. **Architectural comparison of the cgPafA ATP and Pup-binding sites.** Structural depiction of the cgPafA monomer (PDB: 4BJR) illustrating the contrasting geometries of its two functional sites: (**A**) the deep well-defined ATP-binding pocket with ATP bound, and (**B**) the surface-exposed Pup-binding groove with a bound Pup peptide. The key sites are highlighted in yellow. Click to view PafA ATP and Pup^{E} binding sites.

The initial virtual screen for the site identified four compounds—A2, A3, C9, and D2—as top hits, all of which demonstrated favourable docking scores (i.e., > -4.0 kcal/mol; Table 3.4). An identical screening protocol applied to the hybrid mtPafA mtPup-binding region identified a distinct set of four hit compounds: B7, A3, C9, and C11, which also met the binding affinity threshold of -4.0 kcal/mol. Assessment of the species-overlapping compounds A3 and C9 revealed a consistent binding preference for the $mtPup^{IQK}$ site over the $cgPup^{IQK}$ site. This preference was demonstrated by the more favourable docking scores (-6.3 vs. -4.8 kcal/mol for A3; -6.5 vs. -5.5 kcal/mol for C9) and comparable calculated binding free energies (ΔG_{bind} of -37.2 vs. -37.4 kcal/mol for A3; -53.3 vs. -46.05 kcal/mol for C9) (Table 3.5). Furthermore, the two alpha-helical binding pockets (Pupα1 and Pupα2) within the broader Pup-binding groove were screened independently. Screening of this region revealed no appreciable binding by the ligands, suggesting that the region does not serve as a viable binding pocket for the screened *in vitro* ligands in either PafA models.

Table 3.5. Ligand docking scores and MM/GBSA free-binding energies for *in vitro* PafA ligands docked into the Pup binding groove of PafA from *Corynebacterium glutamicum*

Compound Name	Docking score (kcal/mol)	Glide emodel (kcal/mol)	MM/GBSA ΔG_{bind} (kcal/mol)
GGE[a]	**-9.629**	**-47.37**	**-11.16**
GGE[a]	-9.319	-43.825	-12.59
KGGE[a]	-7.345	-52.882	-11.76
KGGE[a]	-6.888	-51.725	-26.65
Compound3_A3[a]	-5.26	-24.029	-22.23
Compound3_A3[a]	-4.936	-20.702	-24.71
Compound33_C9[b]	**-5.542**	**-30.042**	**-53.34**
Compound33_C9[b]	-5.349	-26.423	-22.89
Compound3_A3[b]	-4.777	-25.661	-37.41
Compound3_A3[b]	-4.569	-24.948	-39.93
Compound2_A2[b]	-4.475	-27.689	-47.95
Compound38_D2[b]	-4.106	-47.489	-55.69
Compound2_A2[b]	-4.071	-34.352	-55.46
Compound33_C9[c]	**-6.451**	**-27.645**	**-46.05**
Compoun3_A3[c]	-6.294	-32.45	-37.24
Compound33_C9[c]	-6.072	-29.773	-51.64
Compound3_A3[c]	-6.045	-35.251	-46.42
Compound3_A3[c]	-5.969	-31.141	-35.05
Compound3_A3[c]	-5.758	-34.41	-47.68
Compound3_A3[c]	-5.743	-35.497	-44.04
Compound3_A3[c]	-5.547	-32.183	-27.39
Compound3_A3[c]	-5.448	-28.849	-43.43
Compound35_C11[c]	-5.261	-47.475	-45.3
Compound3_A3[c]	-5.251	-31.642	-46.38
Compound19_B7[c]	-4.395	-34.299	-46.99

[a]docked into cgPupE site, [b]docked into cgPupIQK, docked into mtPup

Docking of the known mtPafA binder ST1926 (Li *et al*., 2022) validated the Pupα1α2 pocket in mtPafA but revealed generally unfavourable binding to cgPafA (-5.16 vs. -3.44 kcal/mol; Table 3.6). Using these results as a baseline, none of the *in vitro* ligands exhibited comparable affinity, suggesting a general lack of appreciable binding to this pocket. A notable exception

was compound A2, which, despite weaker predicted binding, adopted a similar pose to ST1926 within the hydrophobic pocket formed by mtPafA's PHE109, TYR107, THR17, GLN352, and LEU355 residues, hinting at a potential shared binding mechanism. This was hypothesised to be a consequence of the shared adamantyl group between A2 and ST1926. While the Pupα1α2 screen yielded suboptimal results, comparative analysis confirmed the Pup^{IQK} site as the more favourable binding region, with compound C9 demonstrating the strongest predicted affinity.

Table 3.6. Ligand docking scores and MM/GBSA free-binding energies of ST1926 and *in vitro* mtPafA ligands docked into the Pup-binding groove of PafA.

Compound Name	**Docking score (kcal/mol)**	Glide emodel **(kcal/mol)**	MM/GBSA ΔG_{bind} **(kcal/mol)**
ST1926[a]	-3.438	-51.66	-30.80
ST1926[b]	-5.160	-33.21	-42.68
Compound33_C9[c]	**-6.451**	**-27.645**	**-46.05**
Compoun3_A3[c]	-6.294	-32.45	-37.24
Compound33_C9[c]	-6.072	-29.773	-51.64
Compound3_A3[c]	-6.045	-35.251	-46.42
Compound3_A3[c]	-5.969	-31.141	-35.05
Compound3_A3[c]	-5.758	-34.41	-47.68
Compound3_A3[c]	-5.743	-35.497	-44.04
Compound3_A3[c]	-5.547	-32.183	-27.39
Compound3_A3[c]	-5.448	-28.849	-43.43
Compound35_C11[c]	-5.261	-47.475	-45.3
Compoun3_A3[c]	-5.251	-31.642	-46.38
Compound19_B7[c]	-4.395	-34.299	-46.99

[a]cgPupα2 site, [b]mtPupα2 site, [c]mtPupIQK site

Overall, all eight tested compounds demonstrated some capacity for docking into the ATP-binding site of PafA. This site is characterized as a small, deep, and enclosed pocket, consistent with a preference for the smaller ligands known to bind PafA *in vitro* (Figure 3.20A). In contrast, the Pup-binding groove is an open, extended surface that has evolved to accommodate the larger Pup peptide (Figure 3.20B). This architectural difference suggests that favourably targeting sites within this groove would require larger, more extensive ligands. Among the compounds screened, C9 exhibited the most promising binding potential to the $mtPup^{IQK}$ site, warranting its selection for further assessment for the site.

To elucidate the binding dynamics of the top-docking ligands (boldened in Tables 3.4 - 3.6), MD simulations were performed. The initial simulations assessed the overall stability of the

cgPafA system by comparing apo cgPafA with the hybrid mtPafA (mtPup-binding groove) in both the presence and absence of ATP and Mg^{2+}. The root-mean-square deviation (RMSD) calculated over 100 ns revealed that cgPafA, cgPafA-ATP, and cgPafA-Mg^{2+} exhibited similar structural trajectories, with minimal fluctuations between them (differences of 0.10–0.23 Å; Figure 3.21 and Table 3.7).

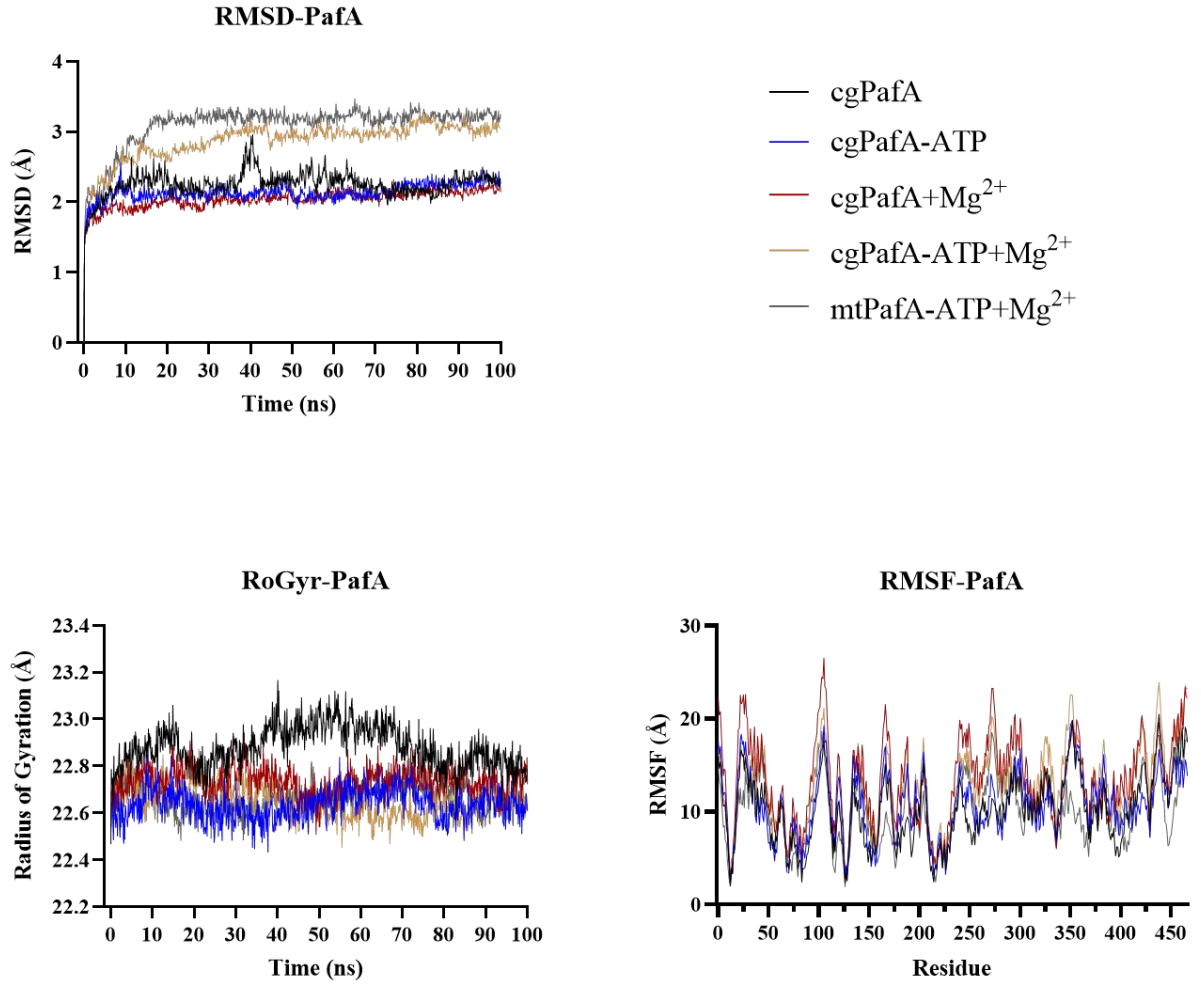

Figure 3.21. Comparative stability of PafA constructs with and without native ligands. Backbone root-mean-square deviation (RMSD-PafA) over simulation time, radius of gyration (RoGyr-PafA) over simulation time, and root-mean-square fluctuation (RMSF) per residue number. Simulations were performed for 100 ns using the Desmond module on the cgPafA monomer (PDB: 4BJR) and a virtual hybrid mtPafA construct, in the presence and absence of ATP and Mg^{2+}.

Apo cgPafA apo, however, exhibited the most variable trajectory, suggesting that both ATP and Mg^{2+} reduce the structural flexibility of the protein over time. This was supported by the radius of gyration (RoGyr) analysis, which showed a greater variation for cgPafA alone (range of 0.70 Å) compared to the other complexes (range of 0.37–0.44 Å) (Figure 3.21, Table 3.7).

The complexes of cgPafA and mtPafA with both ATP and Mg^{2+} displayed similar RMSD profiles but required a longer period to achieve equilibrium (~20 ns) than the other PafA structures (<10 ns). Despite this slower stabilization, their final radius of gyration (RoG) values were comparable to those of the cgPafA-ATP and cgPafA-Mg^{2+} complexes, indicating a similar overall compactness and stability among all bound states. Furthermore, calculated root-mean-square fluctuation (RMSF) values were comparable across all tested structures, with most residues demonstrating similar flexibility profiles (Figure 3.21, Table 3.7).

Subsequent simulations were performed to evaluate the stability of cgPafA in complex with compounds D2, C10, and A3 within the ATP-binding site in the presence of Mg^{2+}, using cgPafA+Mg^{2+} and cgPafA-ATP+Mg^{2+} as controls. All three ligands remained bound throughout the 100 ns simulation, with backbone RMSD and RoG values comparable to the controls, indicating stable complex formation and overall protein fold maintenance (Figure 3.22, Table 3.7).

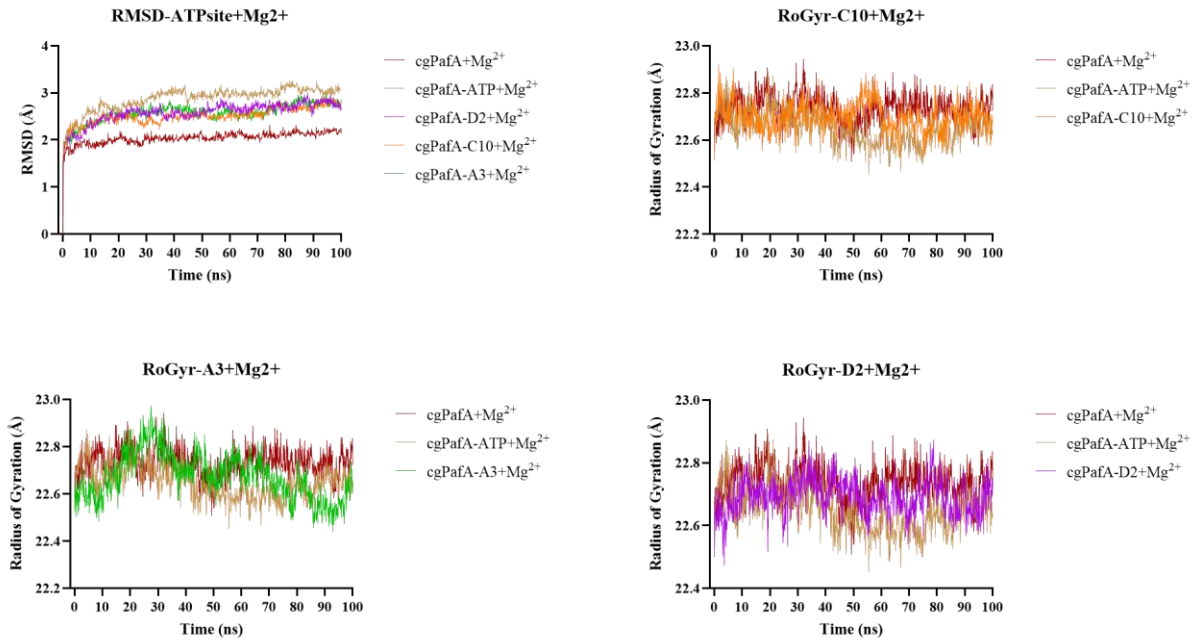

Figure 3.22. Comparative stability of cgPafA construct bound by native and *in vitro* ligands. Backbone RMSD (RMSD-ATPsite+Mg^{2+}) over simulation time, RoGyr of in vitro ligands (RoGyr+A3/C10/D2+Mg^{2+}) over simulation time. Simulations were performed for 100 ns using the Desmond module on the cgPafA monomer (PDB: 4BJR)

Table 3.7. RMSD, Radius of Gyration, RMSF and RMSD of ligand only of cgPafA and ligands docked into the ATP site after 10-100 ns MD simulations

	cgPafA (Mg^{2+})	cgPafA (ATG+Mg^{2+})	cgPafA (D2+Mg^{2+})	cgPafA (C10+Mg^{2+})	cgPafA (A3+Mg^{2+})	cgPafA	cgPafA (ATP)	cgPafA (A3)
RMSF (Å)								
Mean ± SD	2.06 ± 0.9	2.94 ± 0.14	2.64 ± 0.13	2.55 ± 0.13	2.62 ± 0.13	2.29 ± 0.14	2.16 ± 0.10	2.28 ± 0.09
Range	0.51	0.78	0.79	0.71	0.723	1.00	0.56	0.60
Difference (complex-PafA)	0.0	0.88	0.58	0.49	0.56	0.00	0.13	0.01
Radius of gyration (Å)								
Mean ± SD	22.73 ± 0.06	22.64 ± 0.07	22.70 ± 0.06	22.70 ± 0.06	22.68 ± 0.07	22.88 ± 0.09	22.64 ± 0.06	22.80 ± 0.06
Range	0.44	0.44	0.36	0.41	0.54	0.70	0.41	0.38
Difference (complex-PafA)	0.0	-0.09	-0.03	-0.03	-0.05	0.00	-0.24	-0.08
RMSF (Å)								
Mean ± SD	13.45 ± 4.6	12.77 ± 4.2	10.78 ± 3.7	11.03 ± 3.8	12.63 ± 4.4	10.21 ± 3.7	10.75 ± 3.5	15.92 ± 6.0
Range	24.03	21.32	20.34	18.97	22.89	18.97	17.88	28.71
Difference (complex-PafA)	0.0	-0.68	-2.67	-2.24	-0.82	0.00	0.54	5.17
RMSD of ligands in complex (Å)								
Mean ± SD	-	1.69 ± 0.15	1.43 ± 0.42	2.60 ± 0.20	0.93 ± 0.06	-	1.62 ± 0.14	2.02 ± 0.12
Range	-	0.75	1.93	1.58	0.42	-	0.95	0.79

Notably, RMSF analysis revealed that the D2 and C10 complexes induced significantly greater rigidity in the protein compared to the controls and the A3 complex (Figure 3.22, Table 3.7). Ligand-specific RMSD analysis further illustrated that A3 and C10 adopted stable conformations within the binding pocket, exhibiting less fluctuation than D2 (Figure 3.22, Table 3.7). This enhanced stability correlated with a higher number of persistent protein-ligand interactions. The A3 complex was stabilized by a hydrogen bond with ARG418, Pi-cation interactions with TRP440 and HIS211, a salt bridge with GLU16, and a critical metal coordination with the Mg^{2+} ion (Figure 3.23). Comparatively, the C10 complex was characterized by less interactions, including two persistent hydrogen bonds (with ASN132 and THR73) and Pi-Pi stacking with TRP440 (Figure 3.24). In contrast, D2 maintained fewer interactions, primarily a hydrogen bond and Pi-cation interaction with ARG219, despite also coordinating the Mg^{2+} ion (Figure 3.25). This complex was selected for detailed MD simulation.

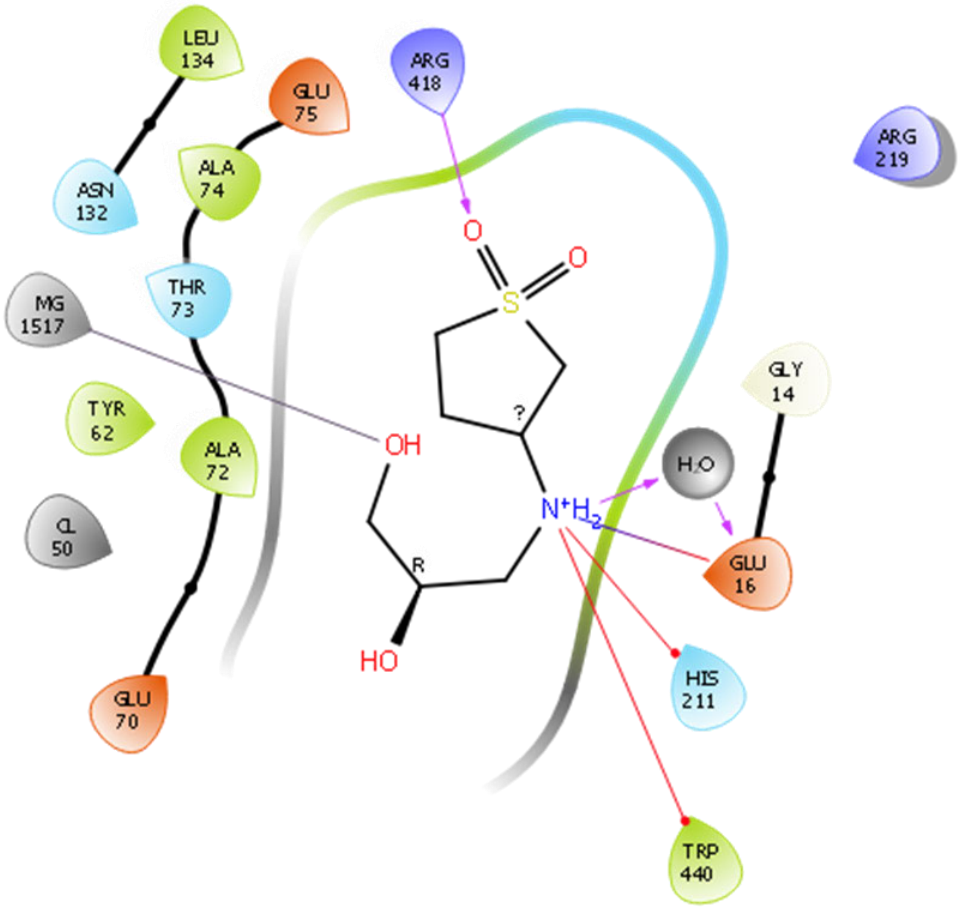

Figure 3.23. **Compound A3 docking cgPafA ATP site**. Docking simulation of compound A3 binding within the deep hydrophobic ATP-binding pocket of cgPafA complexing with a Mg^{2+} ion (purple) (A). A3 interactions post molecular dynamic simulation at 100 ns reveal 5 key interactions of the compound within the site including a stabilizing hydrogen bond with R418, 2X Pi-cation interactions with TRP440 and HIS211, a salt bridge with GLU16, and the critical metal coordination with the Mg^{2+} (B). Click PafA-Mg^{2+}-A3 to view the complex.

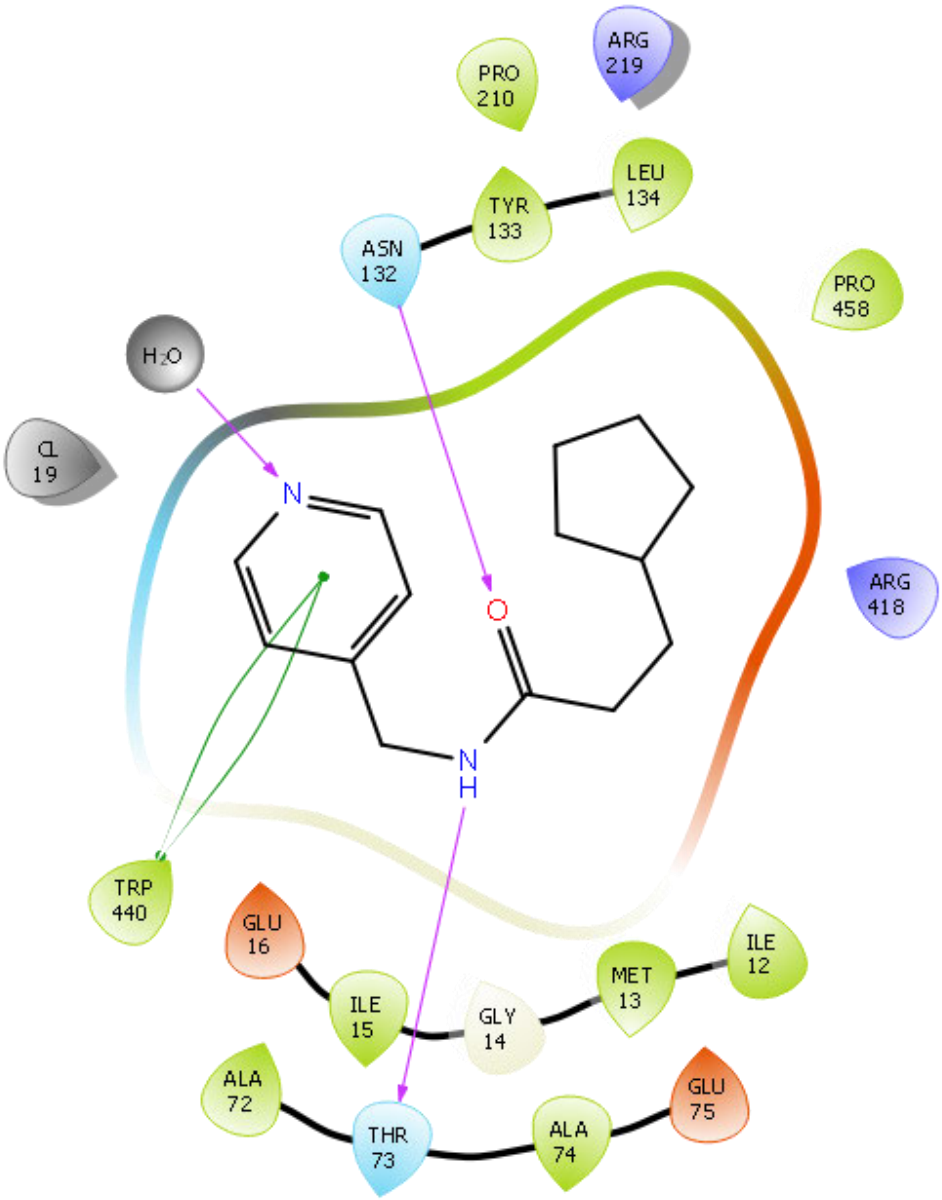

Figure 3.24 Compound C10 docking cgPafA ATP site. Docking simulation of compound C10 binding within the deep hydrophobic ATP-binding pocket of Mg^{2+}-bound cgPafA (A). C10 interactions post molecular dynamic simulation at 100 ns reveal 4 interactions of the compound within the site including 3 hydrogen bonds with N132 and T73 (third hydrogen bond with a water molecule), and a Pi-Pi stacking with W440 (B). Click PafA-Mg^{2+}-C10 to view the complex

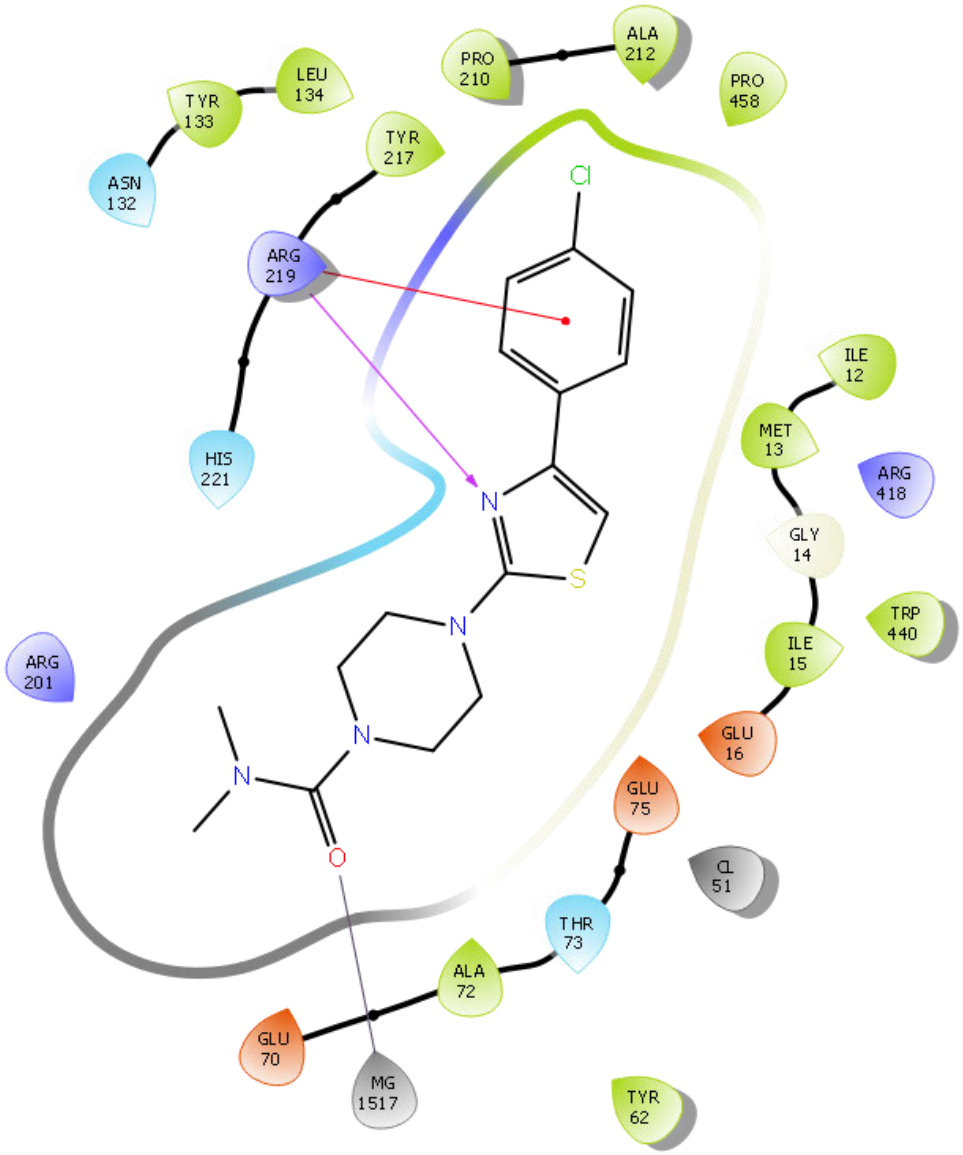

Figure 3.25 Compound D2 docking cgPafA ATP site. Docking simulation of compound D2 binding within the deep hydrophobic ATP-binding pocket of cgPafA complexing with a Mg^{2+} ion (purple) (A). D2 interactions post molecular dynamic simulation at 100 ns reveal only 3 interactions of the compound within the site including a hydrogen bond and Pi-cation interaction with R219, and metal coordination with the bound Mg^{2+} ion (B). Click PafA-Mg^{2+}-D2 to view the complex

To mimic the physiological state in which Pup binding occurs, all simulations were conducted with ATP and Mg^{2+} bound. The simulations confirmed that the hybrid mtPafA structure maintained stable ATP/Mg^{2+} binding, with global metrics (RMSD, RoG, RMSF) comparable to the native cgPafA (Figure 3.21). Structural analysis was corroborated by post-simulation MM/GBSA binding free energy calculations (Table 3.7). Notably, the cgPafA-C10+Mg^{2+} complex demonstrated the most favourable mean total free energy of -47.8 ± 3.3 kcal/mol, which was superior to the native cgPafA-ATP+Mg^{2+} complex (-36.4 ± 8.0 kcal/mol).

In contrast, A3 and D2 yielded unfavourable energies, primarily due to high electrostatic solvation penalties (Table 3.8). Integrating the MD and MM/GBSA results, compound C10 was predicted to be the most promising binder, forming a stable, rigid complex with the most favourable binding affinity within the ATP site of cgPafA.

Table 3.8. Mean free binding energies of cgPafA complexes post molecular dynamics simulations

Compound	$\Delta G_{coulomb}$[a]	$\Delta G_{covalent}$[b]	ΔG_{Hbond}[c]	ΔG_{lipo}[d]	$\Delta G_{packing}$[e]	ΔG_{solvGB}[f]	ΔG_{VdW}[g]	ΔG_{bind}[h]
ATP+Mg^{2+}	-65.50 ± 8.32	0.15 ± 2.67	-6.01 ± 0.56	-9.65 ± 0.53	-3.67 ± 0.62	102.92 ± 6.47	-54.34 ± 5.11	-36.40 ± 7.99
D2+Mg^{2+}	-20.48 ± 2.29	3.10 ± 1.80	0.00 ± 0.03	-22.82 ± 1.19	-5.04 ± 1.05	78.40 ± 2.80	-52.36 ± 2.28	-19.21 ± 4.05
C10+Mg^{2+}	-20.64 ± 3.56	-1.61± 1.00	-0.98 ± 0.15	-13.96 ± 1.23	-2.86 ± 0.89	29.58 ± 2.63	-37.29 ± 3.04	-47.76 ± 3.33
A3+Mg^{2+}	-25.28 ± 4.18	3.25 ± 1.64	-0.99 ± 0.35	-6.83 ± 0.56	0.00 ± 0.00	63.30 ± 2.86	-30.05 ± 2.22	3.39 ± 3.93
ATP	-60.64 ± 5.35	1.59 ± 2.00	-6.61 ± 0.81	-11.20 ± 0.43	-2.41 ± 0.49	41.74 ± 3.31	-50.61 ± 2.46	-88.14 ± 4.44
A3	-39.29 ± 6.38	1.04 ± 0.83	-1.81 ± 0.67	-6.27 ± 0.67	0.00 ± 0.0	43.34 ± 6.49	-31.40 ± 2.57	-34.40 ± 3.89
C9	-4.58 ± 1.18	1.70 ± 0.52	-0.48 ± 0.10	-15.33 ± 0.78	-1.40 ± 0.30	12.39 ± 1.43	-38.64 ± 1.75	-46.35 ± 2.64

[a]Contribution of coulomb energy to the binding free energy, [b]Covalent-bonding contribution to the binding free energy, [c]Hydrogen-bonding contribution to the binding free energy, [d]Lipophilic energy contribution to the binding free energy, [e]Pi-Pi packing energy contribution to the binding free energy, [f]Generalized Born electrostatic solvation energy contribution to the binding free energy, [g]Contribution of Van der Waals interaction energy to the binding free energy, [h]Total binding free energy

Next, to elucidate the specific role of Mg^{2+} in ligand binding, MD simulations were conducted to compare the cgPafA-A3 complex with and without the ion, building on initial docking and MM/GBSA data that suggested A3 was a superior binder in its absence. The global stability of the cgPafA-

A3 complex was similar to the cgPafA-ATP benchmark and superior to the apo protein. Intriguingly, the presence of Mg^{2+} did not enhance this stability; the cgPafA-A3 complex without Mg^{2+} exhibited a lower backbone RMSD, indicating less conformational drift, and a more compact, consistent RoG throughout the simulation (Figures 3.26; Table 3.4). In contrast, RMSF values were substantially elevated in both A3-bound states compared to the apo protein, suggesting that A3 binding induces localized conformational changes regardless of Mg^{2+}.

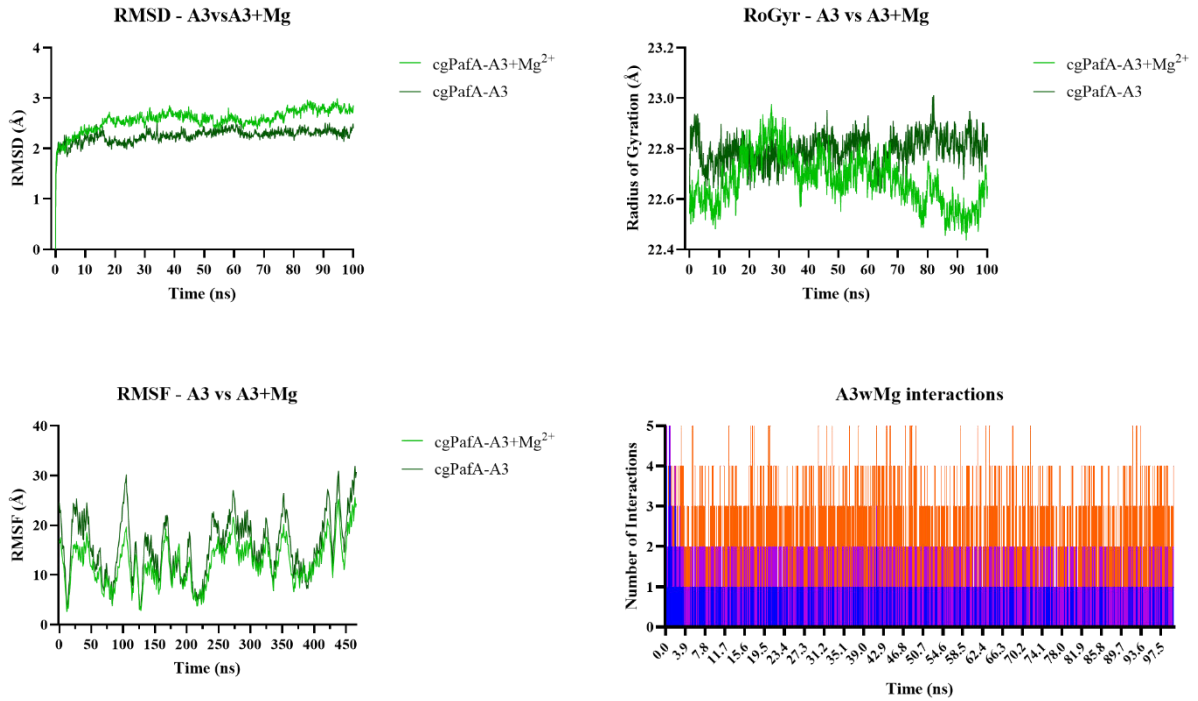

Figure 3.26 Comparative stability of cgPafA construct in complex with ligand A3 with and without Mg^{2+}. Backbone RMSD (RMSD-A3vsA3+Mg^{2+}) over simulation time, RoGyr of A3 complex (RoGyr-A3vsA3+Mg^{2+}) over simulation time, and RMSF (RMSF-A3vsA3+Mg^{2+}) over simulation time. Simulations were performed for 100 ns (RMSF exception, 450 ns) using the Desmond module on the cgPafA monomer (PDB: 4BJR).

Analysis of the ligand itself revealed that A3, once equilibrated, was stable in the ATP pocket in both conditions. However, A3 in the presence of Mg^{2+} showed lower ligand RMSF, correlating with more consistent protein-ligand interactions, including a hydrogen bond with ARG418 and Pi-cation interactions with TRP440 and HIS211 (Figure 3.23). Analysis of the simulation trajectory in the presence of Mg^{2+} revealed that all key interactions, including the salt bridge between A3 and GLU16, remained stable by retaining relatively the same number of respective interactions throughout the entire simulation (a significant drop in the number of hydrogen bonds at the onset of the simulation was, however, noted). This persistence indicated that A3 forms a specific and stable binding pose within the ATP-binding site. In the absence of Mg^{2+}, the A3 hydroxyl group that would otherwise coordinate Mg^{2+} was free to form additional,

dynamic interactions, including hydrogen bonds with water molecules and residues GLU16 and GLU209 (Figure 3.27). This resulted in a higher total number of interactions but with greater variability.

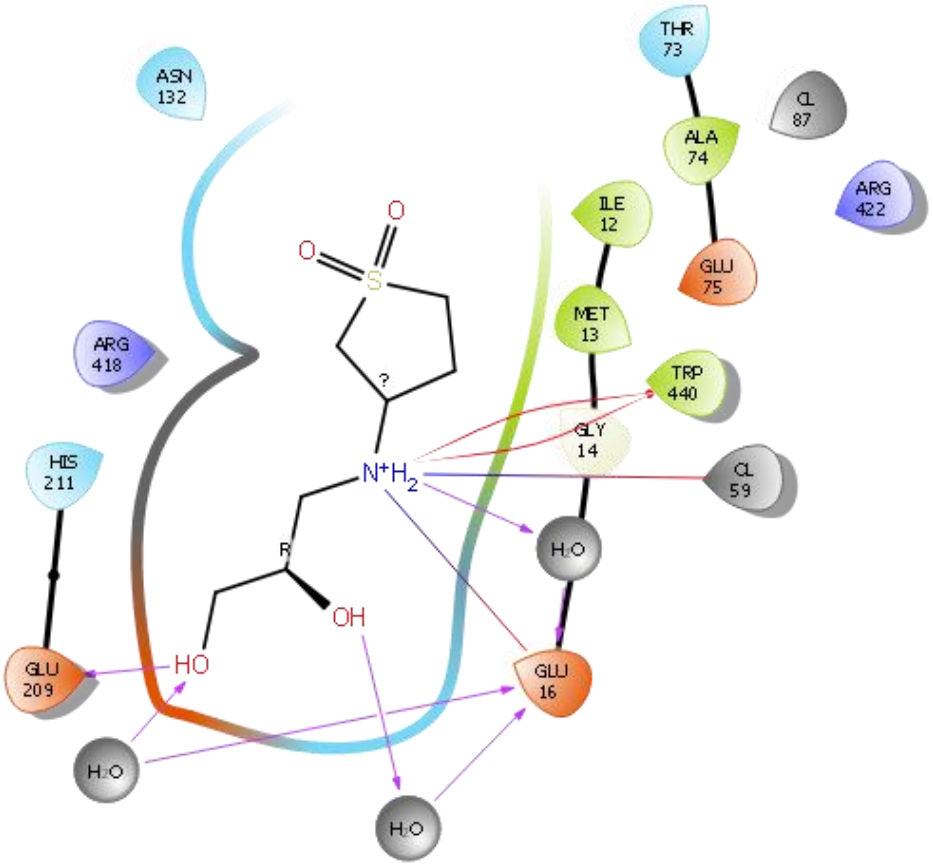

Figure 3.27 Compound A3 docking cgPafA ATP site in the absence of Mg^{2+}. Docking simulation of compound A3 binding within the deep hydrophobic ATP-binding pocket of cgPafA in the absence of a Mg^{2+} ion. A3 interactions post molecular dynamic simulation at 100 ns reveal increased number of dynamic interactions within the site including hydrogen bonds with water molecules and residues E16 and E209. Click PafA-A3 to view the complex

This structural dichotomy was resolved by the binding free energy calculations. Post-MD MM/GBSA revealed a dramatically more favourable total energy for A3 in the absence of Mg^{2+} (-34.4 ± 3.9 kcal/mol) compared to its presence (3.4 ± 4.0 kcal/mol) (Table 3.8). While the A3-Mg^{2+} complex exhibited a more rigid and consistent binding pose, the A3 complex without Mg^{2+} was globally more stable and achieves significantly superior predicted binding affinity. This indicated that the Mg^{2+}-coordinating hydroxyl group of A3 is more beneficial for affinity when solvated and available for dynamic interactions than when locked in a metal coordination bond. Figure 3.28 revealed that Mg^{2+} coordination is critical for binding pose rigidity

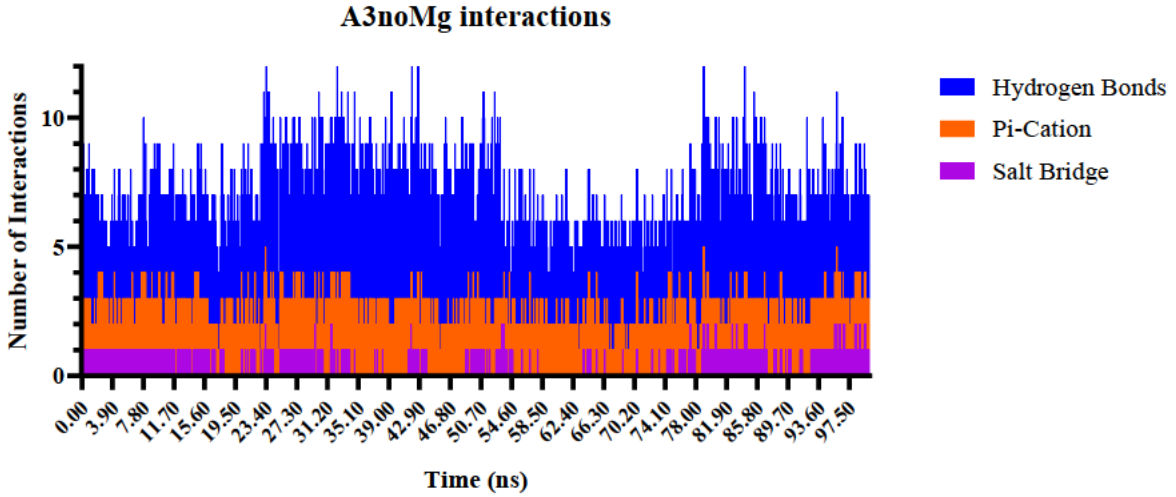

Figure 3.28. Compound A3 interactions within cgPafA ATP site in the absence of Mg2+. MD simulation of compound A3 binding within the deep hydrophobic ATP-binding pocket of cgPafA in the absence of a Mg^{2+} ion revealing key interactions of the compound throughout the simulation.

In the absence of Mg^{2+}, the binding pose of the compound was characterized by a significant increase in the number of hydrogen bonds, although the key pi-cation interaction and salt bridge were still intermittently present. However, these interactions, particularly the hydrogen bonds and salt bridge, exhibited high conformational flexibility throughout the simulation, with the salt bridge breaking and reforming. This increased dynamic variability highlighted the coordinating and stabilizing role of Mg^{2+} in locking the compound into a well-defined pose. Among the eight ligands screened, only compound C9 demonstrated a reliable binding propensity for the Pup-binding groove of both cgPafA and mtPafA, localizing to a specific pocket termed the Pup^{IQK} site (Figure 3.29, Table 3.9).

Compound C9 engaged the binding groove through a combination of polar and aromatic contacts, forming two hydrogen bonds (one water-bridged and one with VAL21) and a stabilizing pi-pi stacking interaction with HIS68. The binding of C9, however, significantly altered the behaviour of the mtPafA complex. The system reached equilibrium more rapidly and exhibited a lower mean backbone RMSD (2.18 Å vs. 3.18 Å for the control) with reduced fluctuation, indicating that C9 stabilizes the protein and reduces its flexibility (Figure 3.30, Table 3.9). This was corroborated by a reduction in residue-level fluctuations (RMSF), consistent with a more rigid and compact structure. C9 itself adopted a stable pose after 30 ns, maintained by persistent interactions including a hydrogen bond with VAL21 and dynamic Pi-Pi stacking with TYR127 and HIS68 (Figure 3.31).

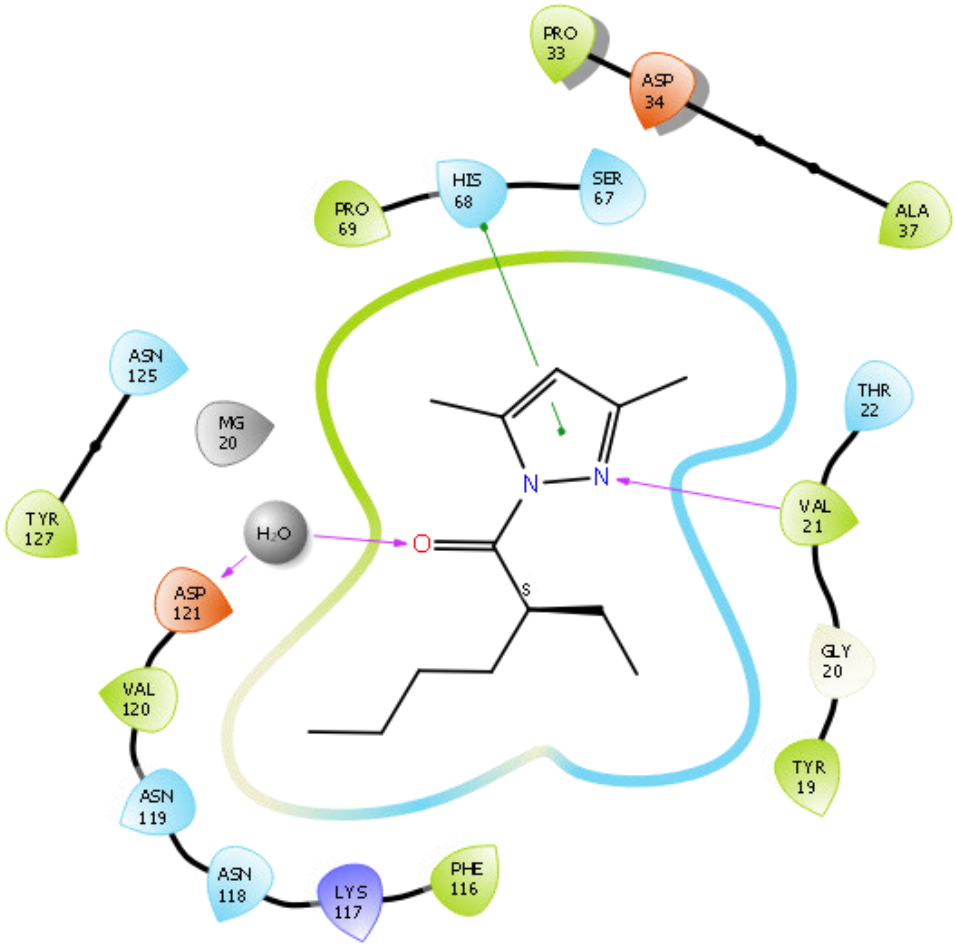

Figure 3.29 Compound C9 docking the mtPafA Pup^{IQK} groove site. Docking simulation of compound C9 binding within the open, surface-exposed Pup^{IQK}-binding groove of mtPafA (A). C9 interactions post molecular dynamic simulation at 100 ns reveal only 3 interactions of the compound within the site, including a hydrogen bond with a water molecule and VAL21, and Pi-Pi stacking interaction with HIS68 (B). Click PafA-C9 to view the complex.

The stability of C9's binding pose within the Pup^{IKQ} site was evidenced by the persistent interaction network observed throughout the simulation (Figure 3.30-3.31). Both hydrogen bonds and the pi-pi stacking interaction remained consistently engaged, confirming a specific and stable binding mode. The ligand's pyrazole methyl group and hydrophobic chain were nestled within complementary hydrophobic pockets. Post-MD MM/GBSA calculations yielded a favourable total free binding energy of -46.35 kcal/mol (Table 3.8), driven by these lipophilic, Pi-Pi, and van der Waals interactions. Collectively, these data suggested C9 binds the Pup^{IQK} site and may exert its effect through an allosteric stabilization mechanism.

Table 3.9. RMSD, Radius of Gyration, RMSF and RMSD of ligand only of mtPafA and mtPafA-C9 complex after 10-100 ns MD simulations

	mtPafA(ATP+Mg2+)	mtPafA(ATP+Mg)-C9
	RMSD (Å)	
Mean ± SD	3.18 ± 0.11	2.18 ± 0.12
Range	0.86	0.62
Difference (complex-PafA)	0.0	-1.00
	Radius of gyration (Å)	
Mean ± SD	22.66 ± 0.06	22.70 ± 0.07
Range	0.37	0.44
Difference (complex-PafA)	0.0	0.04
	RMSF (Å)	
Mean ± SD	10.01 ± 3.5	6.98 ± 2.4
Range	16.75	12.73
Difference (complex-PafA)	0.0	-3.13
	RMSD of C9 in complex (Å)	
Mean ± SD	-	2.49 ± 0.19
Range	-	1.27

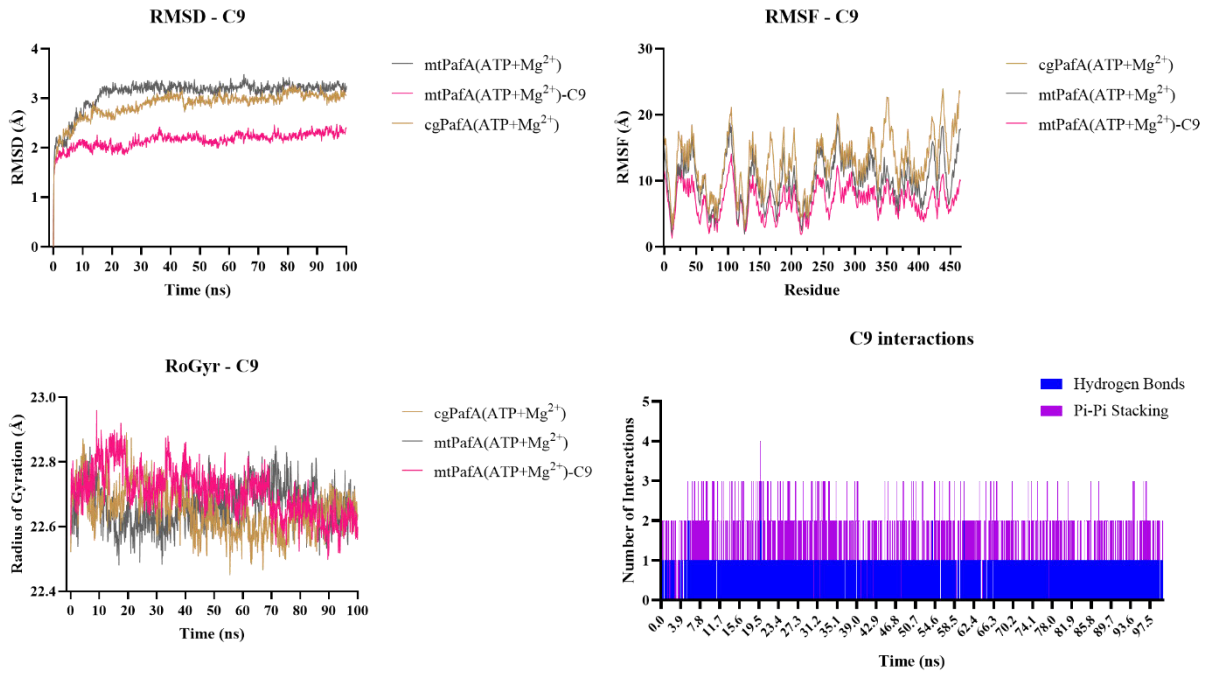

Figure 3.30 Comparative stability of cgPafA construct in complex with ligand C9. Backbone RMSD (RMSD-C9) over simulation time, RoGyr of C9 comlex (RoGyr-9) over simulation time, and RMSF (RMSF-C9) over simulation time. Simulations were performed for 100 ns using the Desmond module on the cgPafA monomer (PDB: 4BJR) that was virtually mutated into mtPafA.

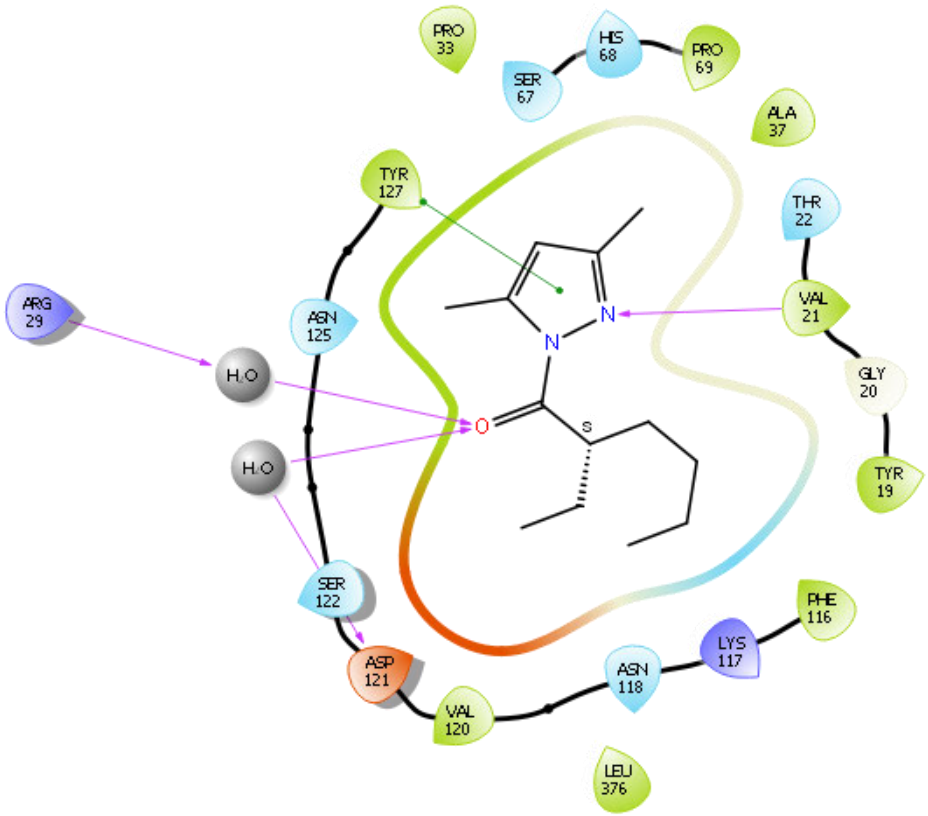

Figure 3.31 Compound C9 docking the mtPafA Pup^{IQK} groove site. Docking simulation of compound C9 binding within the open, surface-exposed Pup^{IQK}-binding groove of mtPafA (A). C9 interactions during molecular dynamic simulation at 40 ns reveal 4 interactions of the compound within the site, including a hydrogen bond with two water molecules and V21, and Pi-Pi stacking interaction with W127 (B).

3.2.3 PafA ligands validation via the cellular thermal shift assay (CETSA)

3.2.3.1 PafA CETSA optimization

To corroborate the *in silico* and TSA findings, the CETSA was employed to evaluate the binding of the identified compounds to PafA in a cellular context. In this method, a change in protein thermal stability (aggregation temperature, Tagg) similarly indicates compound binding. To ensure consistent analysis, equal protein loading for CETSA samples transferred to nitrocellulose membranes was verified using Ponceau S staining. This technique relied on the consistent presence of a thermostable protein across the CETSA working temperature range. A representative blot demonstrating efficient transfer and equal loading (up to 59 °C) is shown (Figure 3.32).

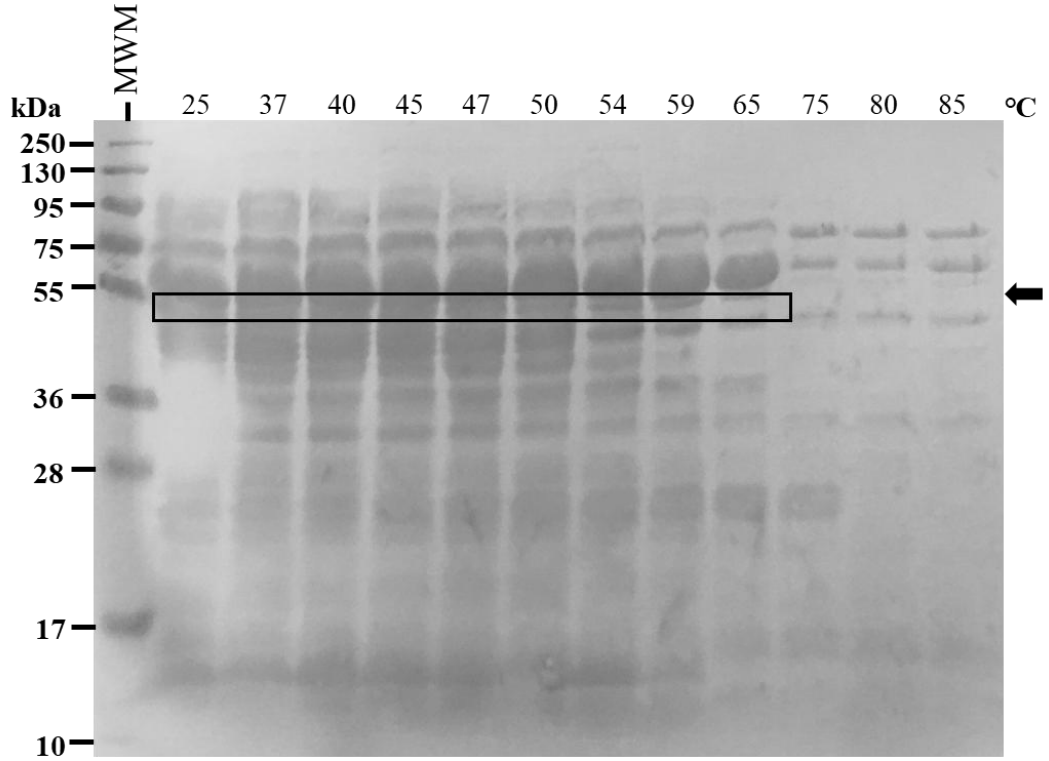

Figure 3.32 Ponceau S Stain of PafA CETSA samples. Soluble protein fractions from the PafA CETSA were separated by 10% SDS-PAGE and transferred to a 0.2 µm nitrocellulose membrane via semi-dry transfer (10 V, 1 h). The membrane was stained with Ponceau S to confirm uniform transfer efficiency and total protein loading using a thermostable protein before probing for PafA. The band corresponding to the expected molecular weight of PafA is highlighted.

Ponceau staining of the blotted membranes confirmed efficient transfer of the full molecular weight range of soluble cellular proteins, from >10 kDa to ~95 kDa, which encompasses the ~54 kDa size of PafA (Figure 3.32). The staining also served as a critical loading control by revealing the thermal stability profiles of endogenous proteins. Notably, a band at ~58 kDa, identified as the chaperone GroEL (including both endogenous and recombinantly co-expressed

with PafA), demonstrated stability up to 59 °C, consistent with its known role as a heat-shock protein. A second stable protein at ~73 kDa, which co-purifies with PafA and presumed to be the putative DnaK chaperone, was also observed.

The inherent stability of GroEL was subsequently used as a loading control for the PafA CETSA experiments (Delport and Hewer, 2022), monitoring its consistent expression across the stable temperature range of 25–65 °C. The PafA CETSA was therefore employed to analyse the thermal stability of PafA in a cellular context. Immunoblot analysis and subsequent densitometry quantification generated PafA melt curves, enabling quantification of PafA stability upon ligand binding (Figure 3.33).

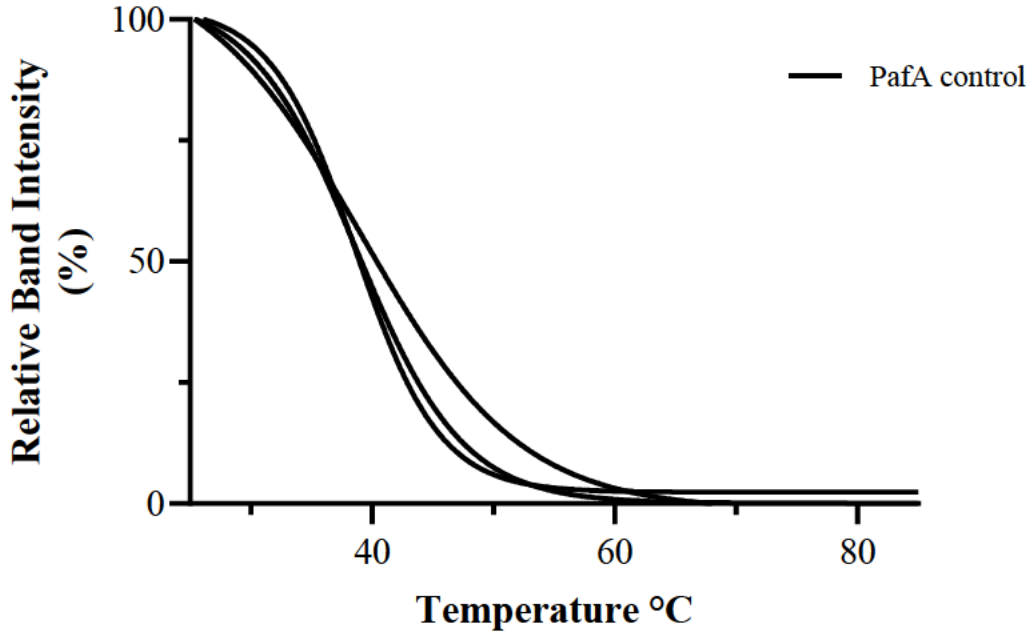

Figure 3.33. Recombinant PafA cellular thermal shift assay (CETSA) optimisation. Recombinant PafA lysates were incubated at a gradient of temperatures. After centrifugation (20,000 × g, 10 min), soluble fractions were resolved by 10% SDS-PAGE and transferred to 0.2 µM nitrocellulose. PafA was detected by immunoblotting using a mouse monoclonal anti-6xHis primary antibody and a goat anti-mouse monoclonal HRP-conjugated secondary antibody, followed by enhanced chemiluminescence. Band intensity was quantified by densitometry, and the resulting data fitted by non-linear regression to calculate the protein aggregation temperature (Tagg) (n=3).

The analysis of the PafA thermal melt curve established a baseline calculated PafA Tagg of 38.80 ± 0.15 °C. A sharp, cooperative thermal denaturation profile of PafA was observed between 25–40°C (Figure 3.33), suggesting a high degree of conformational sensitivity for the protein within the cellular milieu. This pronounced transition provided an excellent signal-to-noise ratio for the PafA CETSA, making it highly responsive to ligand binding.

With the baseline assay optimized, its functionality was validated using the control ligands bithionol and ST1926. As anticipated, both compounds induced a stabilising effect on PafA relative to the control (Figure 3.34).

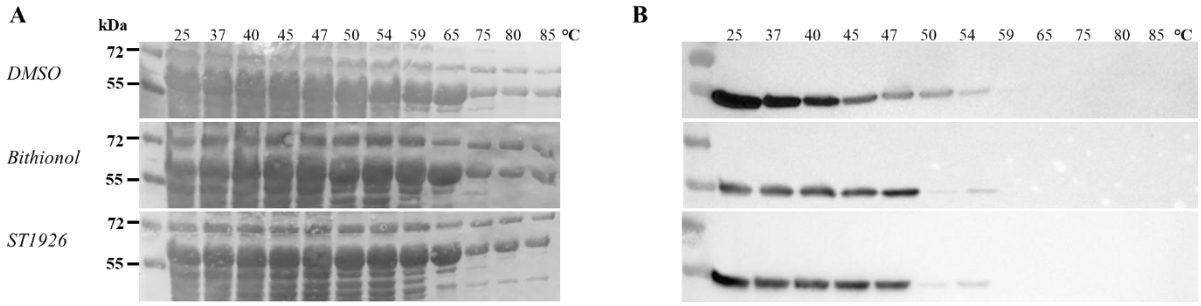

Figure 3.34. Control PafA CETSA. PafA lysates were exposed overnight to bithionol, ST1926, or DMSO, then subjected to the CETSA protocol (thermal gradient, centrifugation at 20,000 × g for 10 min). The soluble fractions were resolved by 10% SDS-PAGE and transferred to a 0.2 µm nitrocellulose membrane, and a (A) Representative Ponceau S stain confirming uniform transfer and loading using GroEL was carried out. (B) Corresponding immunoblots for PafA, detected using a mouse anti-6xHis primary antibody, an HRP-conjugated goat anti-mouse secondary antibody, and enhanced chemiluminescence (ECL) (n=3).

The PafA CETSA revealed a pronounced ligand-induced stabilization of PafA. Qualitatively, immunoblots revealed that both compound-exposed PafA samples retained solubility across a wider temperature range (25-47°C), in stark contrast to the rapid decline observed in the no-ligand control within the same range (Figure 3.34B). This visual assessment was confirmed and quantified by densitometry analysis. The resulting melt curve for bithionol-bound PafA displayed a prototypical sigmoidal shape with a gentle slope and an initial plateau, confirming the presence of folded, soluble protein. In contrast, the control curve decreased sharply at lower temperatures. Quantification of these curves demonstrated that bithionol binding significantly increased the PafA Tagg (ΔTagg) by 6.18°C, increasing from 38.80 ± 0.15°C to 47.28 ± 1.43°C (Figure 3.35B, Table 3.10).

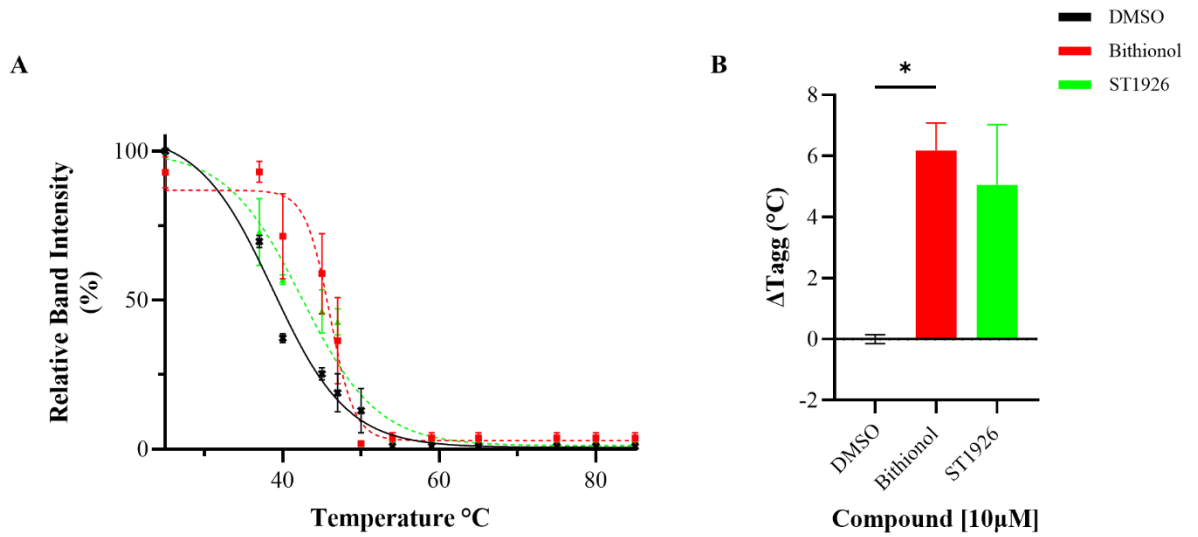

Figure 3.35. Recombinant PafA ΔT_{agg} in the presence of control ligands. The optimized PafA CETSA was used to assess thermal stability in the presence of 10 µM bithionol or ST1926. (A) Thermal melt curves derived from quantitative western blot analysis. (B) Calculated Tagg for each treatment compared to the DMSO control. Data are mean Tagg ± SEM (n=3); statistical significance was determined by one-way ANOVA with multiple comparisons (*p<0.05).

The significant increase in PafA Tagg upon bithionol binding was consistent with ligand stabilization, wherein the compound selectively binds and populates the folded state of the protein, shifting the equilibrium toward the soluble form and raising the aggregation temperature. The stabilizing effect was not unique to bithionol. The compound ST1926 also induced a substantial thermal shift, increasing the PafA Tagg by 5.05 °C from 38.80 ± 0.15 °C to 47.27 ± 0.62 °C. This increase represented a strong trend, closely approaching statistical significance (p = 0.0512) (Figure 3.34-3.35, Table 3.10). The nearly identical stabilization by both ligands confirmed PafA as a readily ligandable target that can be monitored by the optimised PafA CETSA. Notably, the strong cellular stabilization conferred by both ligands directly contrasted with their destabilizing effect in the purified TSA system, where they populated a less stable PafA conformation. This reversal was likely attributable to the more physiologically complex and protective environment of the cell. This functional divergence underscores the complementary nature of the two techniques: TSA probing protein unfolding directly, while CETSA reports on aggregation of the denatured state, together providing a comprehensive profile of ligand-induced conformational modulation

Table 3.10. PafA CETSA analysis in the presence of the identified PafA ligands

Compound	Tagg (°C)	Tagg shift	P-value	Significance (*P < 0.05*)
DMSO control	38.80 ± 0.15	-	-	-
Bithional	47.28 ± 1.43	6.18	0.0233	Yes
ST1926	47.27 ± 0.62	5.05	0.0512	No

Statistical analysis via One-way ANOVA multiple comparisons testing (Tagg as mean Tagg ± standard error; n=3)

Having established a functional 'PafA CETSA' assay—validated by the potent stabilization induced by control ligands bithionol and ST1926, thereby confirming the assay's suitability for profiling the cellular target engagement by ligands—the validated assay was subsequently applied to translate the *in vitro* TSA findings into a more biologically relevant context**.**

3.2.3.1 PafA CETSA: small-molecule ligands validation

To evaluate cellular target engagement, the seven novel PafA ligands identified *in vitro* were subjected to the validated PafA CETSA protocol. The thermal stability of compound-treated PafA samples was similarly analysed by immunoblotting across a temperature gradient, and a representative Ponceau S stain was done to confirm consistent total protein loading using GroEL as a loading control (Figure 3.36).

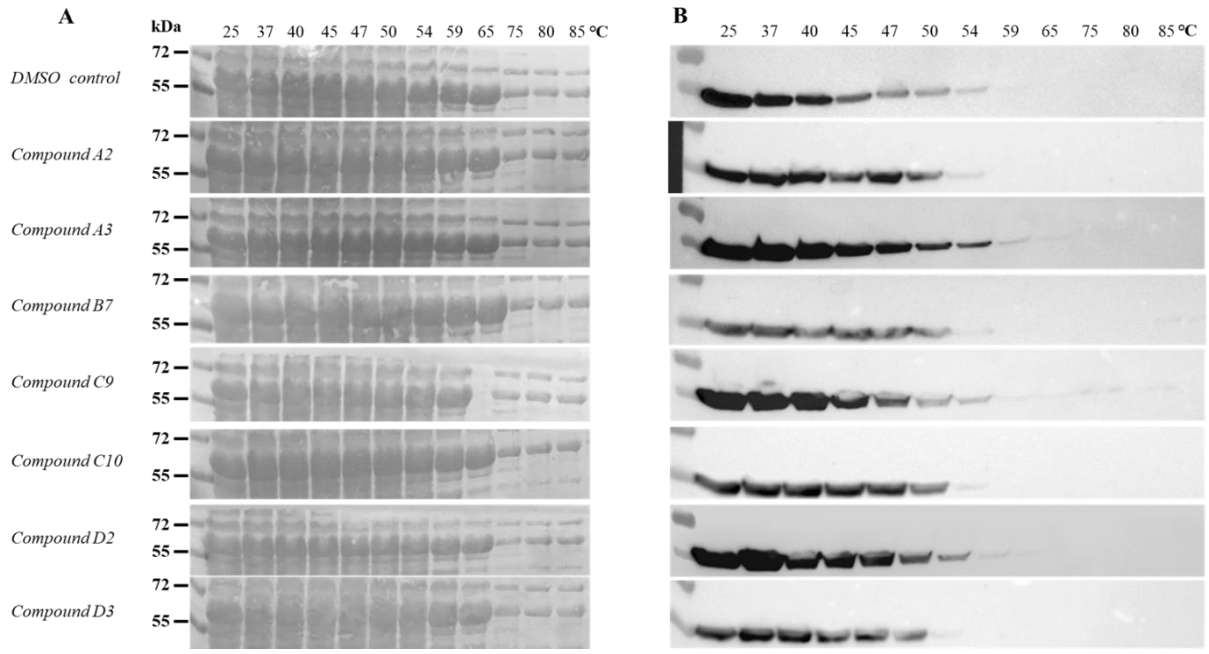

Figure 3.36. PafA cellular thermal shift assay (CETSA) in the presence of the identified PafA binders. PafA samples exposed overnight to the respective compounds were incubated at increasing temperatures, centrifuged (20 000 x *g*, 10 min), and the soluble fractions separated by 10% SDS-PAGE and subsequently semi-dry transferred onto 0.2 µM nitrocellulose. Ponceau staining (**A**) to assess transfer efficiency was done following each transfer and subsequent PafA immunoblot detection (**B**) of the respective transferred samples done using mouse monoclonal anti-6x His antibodies and goat anti-mouse secondary antibodies and enhanced chemiluminescence substrate. Densitometry analysis for each immunoblot treatment was conducted and the data subsequently used to calculate PafA Tagg via non-linear regression modelling (n=3).

All seven novel compounds (A2, A3, B7, C9, C10, D2, D3) from the TSA screen reproduced the stabilization effect previously observed with the control ligands. Qualitative immunoblot analysis revealed a clear increase in PafA solubility across the initial 25-47°C range for all the hit compounds compared to the control (Figure 3.36B). The qualitative analysis was confirmed by quantitative densitometry. The resulting thermal melt curves for all seven compounds exhibited the prototypical right shift, with a pronounced initial plateau that indicated the presence of a stable population of folded, soluble protein (Figure 3.37A). PafA Tagg calculation from the melt curves definitively provided corroborative quantitative data, demonstrating a statistically significant increase for every compound, thereby validating them as bona fide

cellular stabilizers of *Mtb* PafA (Figure 3.37B, Table 3.11). The stabilization of PafA by all seven compounds was both dramatic and statistically robust, with each compound inducing a Tagg increase significantly greater by more than 7 °C ($p \leq 0.0002$).

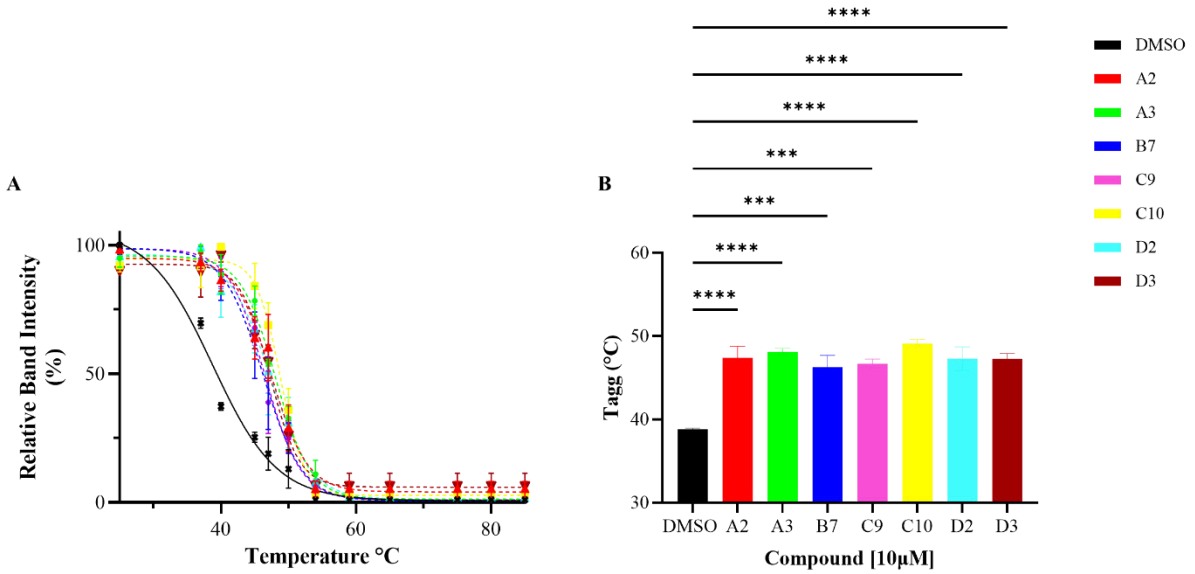

Figure 3.37. Thermal stabilization of PafA by small-molecule drug-like compounds. The optimized PafA CETSA was used to (**A**) profile the thermal stability of PafA in the presence of 10 µM compounds from the ChemBridge DiverSET library (A2, A3, B7, C9, C10, D2, D3). Calculated Tagg values (mean Tagg ± standard error, n=3) for each compound were statistically compared to the DMSO control via one-way ANOVA with multiple comparisons testing (***p<0.001; ****p<0.0001)

Table 3.11. PafA CETSA analysis in the presence of the identified PafA ligands

Compound	Tagg (°C)	Tagg shift	P-value	Significance (*P < 0.05*)
DMSO control	38.80 ± 0.15	-	-	-
A2	47.40 ± 1.36	8.60	<0.0001	Yes
A3	48.11 ± 0.405	9.31	<0.0001	Yes
B7	46.26 ± 1.43	7.46	0.0002	Yes
C9	46.65 ± 0.59	7.85	0.0001	Yes
C10	49.09 ± 0.50	10.29	<0.0001	Yes
D2	47.28 ± 1.43	8.48	<0.0001	Yes
D3	47.27 ± 0.62	8.47	<0.0001	Yes

Statistical analysis via One-way ANOVA multiple comparisons testing (Tagg as mean Tagg ± standard error; n=3)

3.3. Effect of *In vitro* ligands in the mycobacterial PPS

3.3.1. *In vitro* ligands effect on mycobacterial proliferation

The physiological impact of the candidate *in vitro* ligands on mycobacterial proliferation was first assessed under standard, non-stressed growth conditions (Figure 3.38). This baseline assessment aimed to distinguish general antibacterial effects from specific inhibition of the inducible PPS stress response. To validate the sensitivity of the assay, cells were treated with kanamycin, which resulted in complete growth inhibition, as expected (Figure 3.38, Table 3.12). Similarly, treatment with the positive control ST1926 significantly suppressed proliferation, consistent with its reported low-micromolar activity against *Mtb* (Li *et al*., 2022).

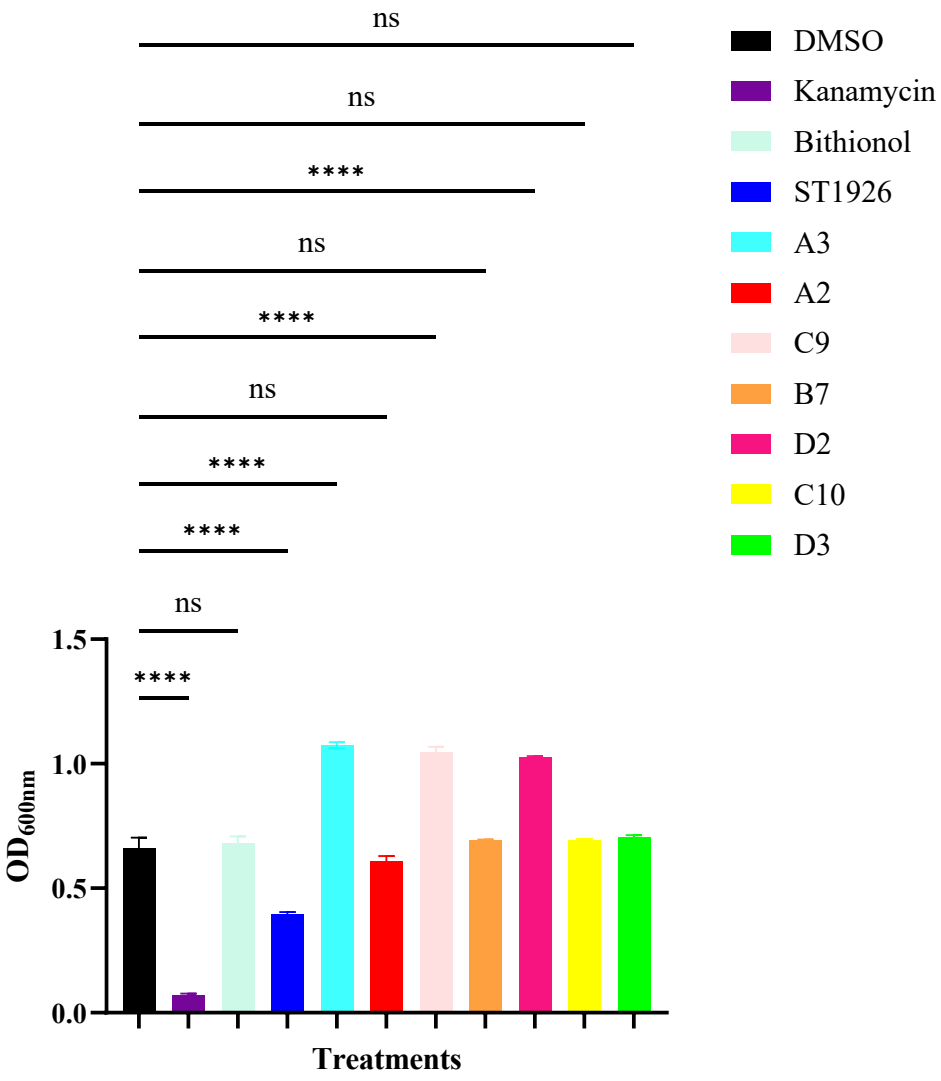

Figure 3.38. PafA *in vitro* ligands effect on mycobacterial cell proliferation. Candidate ligands effect on *Msm* mc^2155 cell growth was conducted by exposing cell culture to ligands at 10 µM final concentrations for 24 h at 37 °C. Control ligands ST1926 (10 µM), bithionol (50 µM), and kanamycin (10 µg/mL) were similarly exposed to their respective cultures, and the growth analysis data were analysed by One-way ANOVA multiple comparisons testing (n=3; **** $p<0.0001$).

In contrast, bithionol demonstrated no anti-proliferative effect even at 50 μM, which diverges from previously reported mycobacterial inhibition (Li *et al*., 2022). Of the seven candidate PafA ligands, only three—A3, D2, and C9—significantly altered *Msm* growth under standard conditions (Figure 3.38, Table 3.12). Unexpectedly, however, these compounds did not inhibit growth but rather induced a significant increase in proliferation compared to the untreated control.

Table 3.12. PafA *in vitro* ligands effect on *Msm* cell proliferation under standard conditions.

Compound	Growth (OD_{600nm})	ΔCell growth (%)	P-value	Significance ($P < 0.05$)
DMSO control	0.662 ± 0.040	-	-	-
Kanamycin	0.071 ± 0.006	-89.27	<0.0001	Yes
Bithionol	0.680 ± 0.029	2.72	0.9933	No
ST1926	0.398 ± 0.006	-39.88	<0.0001	Yes
A2	0.611 ± 0.018	-7.70	0.3237	No
A3	1.074 ± 0.012	62.24	<0.0001	Yes
B7	0.692 ± 0.003	4.53	0.8558	No
C9	1.074 ± 0.012	62.24	<0.0001	Yes
C10	0.694 ± 0.004	4.83	0.8025	No
D2	1.029 ± 0.002	55.44	<0.0001	Yes
D3	0.704 ± 0.009	6.34	0.5557	No

Statistical analysis via One-way ANOVA multiple comparisons testing (growth as mean OD_{600nm} ± standard error; n=3)

3.3.2. *In vitro* ligand effects under oxidative stress growth conditions

Prior to ligand testing, a sub-lethal $NaNO_2$ concentration was established to sensitize the proliferation assay to PPS inhibition. This was achieved by treating *Msm* with a $NaNO_2$ dilution series to identify a concentration that perturbs but does not arrest growth (Figure 3.39). Exposure of the cell culture to $NaNO_2$ induced a dose-dependent decrease in cell proliferation, the cells exhibiting a high degree of sensitivity to nitrosative stress—viability decreasing by 32% at 1.25 mM and by 72% at 2.5 mM $NaNO_2$.

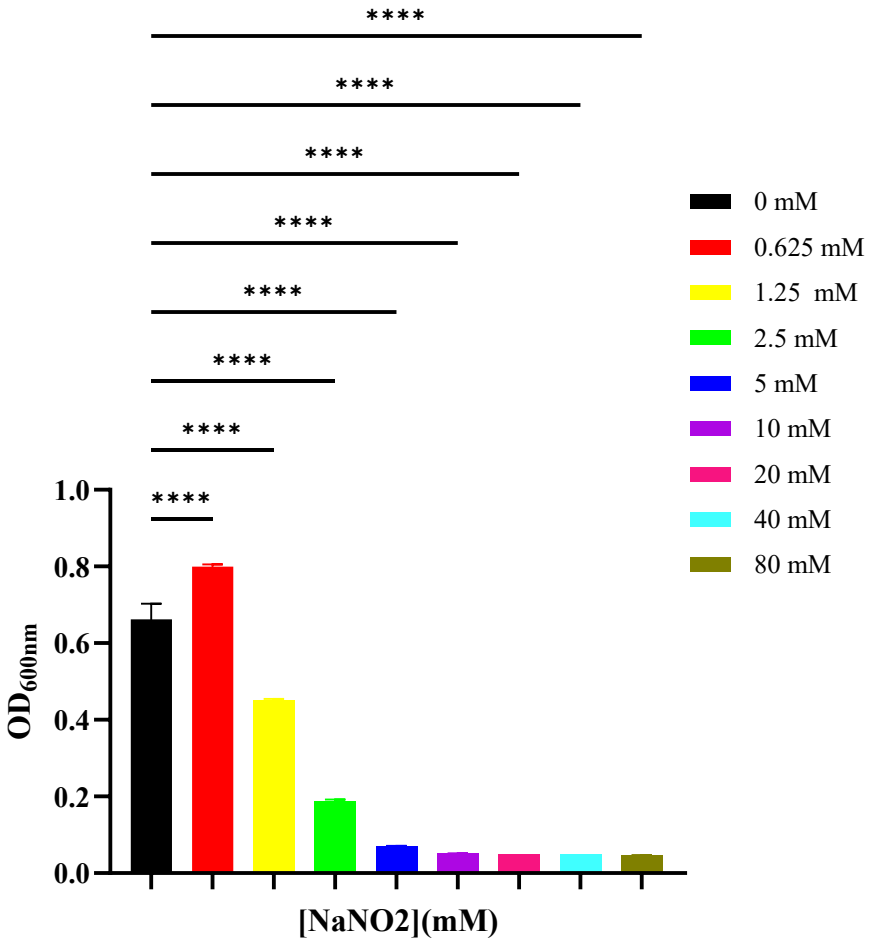

Figure 3.39. Nitrosative stress mycobacterial proliferation assay optimisation. Increasing $NaNO_2$ concentrations were respectively exposed to prepared *Msm* culture for 24 h at 37 °C. Growth analysis data was subsequently analysed by One-way ANOVA multiple comparisons testing (n=3; **** $p<0.0001$).

Conversely, at the lowest tested concentration (0.625 mM), a significant increase in proliferation was observed, suggesting a potential hormetic response. Based on this profile, a concentration of 2.5 mM $NaNO_2$ was selected for subsequent assays. This concentration induced a substantial metabolic stress on cells without causing complete growth arrest, thereby

potentially sensitizing the system to detect modulation of the PPS-dependent stress response. Compound evaluation at this oxidative stress concentration was subsequently conducted as outlined below (Figure 3.40). In the nitrosative stress assay, the ligands did not exhibit significant alteration of *Msm* proliferation relative to the stressed control. Growth data for cells treated with each ligand in the presence of 2.5 mM $NaNO_2$ were similar to those for the stress-only control. This lack of effect was comparable to the reference compounds ST1926 and bithionol under the same conditions. Detailed proliferation data for ligands A3, C9, and D2 are not shown, but followed this trend.

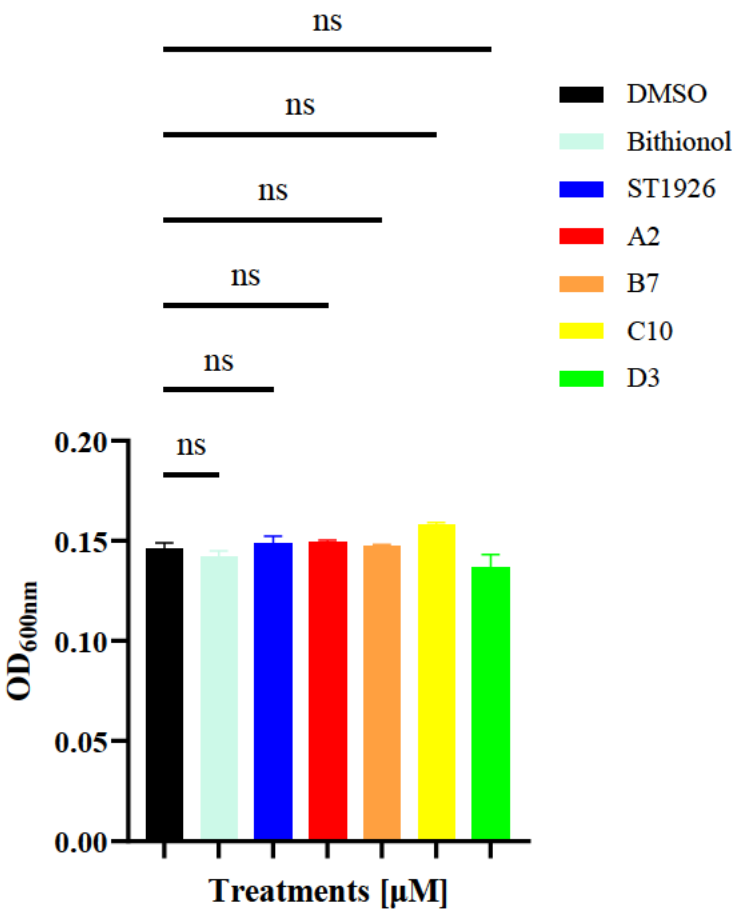

Figure 3.40. PafA in vitro ligands effect on mycobacterial cell proliferation under oxidative stress conditions. Candidate ligands effect on *Msm* mc^2155 cell growth exposed to 2.5 mM $NaNO_3$ oxidative stress was conducted by incubating cell culture with ligands at 10 µM final concentrations for 24 h at 37 °C. Control ligands ST1926 (10 µM), bithionol (50 µM) were similarly exposed to their respective cultures, and the growth analysis data were analysed by One-way ANOVA multiple comparisons testing (n=3).

CHAPTER 4. DISCUSSION

4.1 *Mtb* PafA is a constitutively active monomer, streamlining pupylation activation

A thorough understanding of a target's structure and mechanism is a prerequisite for rational drug discovery. This is particularly critical for *Mtb* PafA, the terminal enzyme in the bacterial pupylation cascade and the sole ligase responsible for conjugating Pup^E to protein substrates, thereby targeting them for proteasomal degradation. The mechanism of this pupylation reaction has been comprehensively delineated, revealing that PafA catalyzes a two-step transfer: first, the phosphorylation of the C-terminal glutamate γ-carbon of Pup^E, followed by the ligation of the activated Pup^E to a specific lysine residue on the target protein. This detailed mechanistic understanding provides a structural and functional framework within which to interpret the inhibitory / ligand potential of the small drug-like molecules identified in this study (Sutter *et al.*, 2010; Guth *et al.*, 2011).

The functional oligomeric state of PafA is a critical consideration for inhibitor screening, as evidence pertaining to the characterised homologs suggests the monomer is the catalytically active species, while the domain-swapped dimer represents a dormant form. This was highlighted by Plourde *et al*. (2025), who demonstrated that recombinant *Cgl* PafA exists in a dynamic equilibrium between monomeric and dimeric states, with the only available X-ray structure for this homologue crystallised in its dimeric state. A similar dimer-based model for *Msm* PafA was also proposed by Ofer *et al.* (2013), based on mathematical modelling, suggesting protein regulation via dimer dormancy may be conserved across multiple actinobacterial homologs. To ensure the relevance of the screening results, it was therefore imperative to elucidate the state of *Mtb* PafA in solution. To this end, this study optimised the production of soluble, recombinant *Mtb* PafA in *E. coli* using an autoinduction system co-expressing the GroEL-GroES chaperonin complex. This chaperone-assisted strategy was essential to overcome the well-documented insolubility of wild-type PafA, which often necessitates solubility-enhancing mutations (Li *et al*., 2022) that may inadvertently alter the protein's native conformation and activity.

Recombinant expression of *Mtb* PafA, assisted by the GroEL-GroES chaperonin system, led to overproduction of a protein at the expected molecular weight (54 kDa) for the monomer by SDS-PAGE analysis. While SDS-PAGE and immunoblotting confirmed successful expression,

the technique was insufficient to determine the protein's native oligomeric state, as the denaturing conditions disrupt non-covalent interactions that would stabilize domain-swapped dimers reported for the *Cgl* homologue (Barandun *et al.*, 2012)—assuming a similar mechanism exists for *Mtb* PafA. To probe the solution-state conformation, purified *Mtb* PafA was analysed using non-reducing SDS-PAGE. This revealed two closely migrating bands, initially suggestive of distinct monomeric species. However, subsequent experiments determined that this heterogeneity was an artifact of incomplete disulfide bridge reduction under the low-beta-mercaptoethanol conditions, rather than evidence of a domain-swapped dimer. This was confirmed by the partial collapse of one of the PafA populations upon the addition of 1 mM DTT prior to analysis under "non-reducing" conditions, indicating the presence of a reversibly oxidized monomeric form.

The purified protein's high sensitivity to oxidation necessitated storage in a reducing solution to maintain stability. This was quantitatively corroborated by the optimised PafA thermal melt assay, where the addition of DTT significantly enhanced the stability of the protein by 7.26 °C. Conversely, in the absence of a reducing agent, the thermal melt profile indicated that PafA adopts a partially unfolded conformation in solution. This state is characterized by the exposure of hydrophobic residues, which accounts for the elevated initial fluorescence signal. This confirmed the intrinsic instability of the protein under non-reducing conditions and suggests that disulfide bonding or a reducing environment is critical for maintaining the proteins native, soluble fold. Critically, the absence of any higher molecular weight species corresponding to a stable dimer provided indirect evidence that the purified protein exists predominantly as a monomer in solution, consistent with the intended goal of screening the catalytically active form. Importantly, however, the treatment composition during non-reducing SDS-PAGE consists of 1% (w/v) SDS, which would disrupt the non-covalent interactions formed by domain-switched dimers. To definitively establish that purified *Mtb* PafA exists exclusively as a monomer in solution, purified protein was analyzed by molecular-exclusion chromatography and native PAGE following Ni^{2+}-NTA purification. SEC on a Sephadex G75 matrix (separation range: 3–80 kDa) allowed for the resolution of monomers from higher-order oligomers. Since a stable dimer or larger oligomer (≥108 kDa) would be excluded from the matrix and elute in the void volume, the complete absence of PafA in this fraction ruled out such species. The protein instead eluted after the 73 kDa contaminant DnaK, yielding an apparent molecular weight consistent with its monomeric form. The absence of higher molecular weight species provided direct biophysical evidence that the recombinant *Mtb* PafA, under the specified

purification conditions, exists predominantly as a monomer in solution. This conclusion was complemented by native PAGE immunoblot analysis, which detected only a single species following Ni^{2+}-NTA purification, indicating the homogenous existence of *Mtb* PafA in solution. Moreover, analysis of the *Mtb* PafA thermal melt profile further supported the conclusion that Mtb PafA is monomeric. The observed single-phase cooperative unfolding transition is characteristic of a monomeric protein, in contrast to the multiphasic denaturation profiles typically exhibited by stable oligomers (Burgos *et al.*, 2008).

The conclusion that *Mtb* PafA is exclusively monomeric may suggest a pathogen-specific adaptation of the PPS enzyme-level regulation. While PafA homologs in non-pathogenic actinobacteria are proposed to be dimeric, regulated by a dormancy mechanism (Brandun *et al*., 2012; Ofer *et al*., 2013; Plourde *et al*., 2025), *Mtb* may have evolved a constitutively active monomeric form. This divergence could reflect the distinct physiological demands of pathogenesis, where instead of intermittent activation via dimer dissociation, *Mtb* requires sustained PPS activity to manage the continuous stresses of host infection, favouring a perpetually active monomer. As support, while all the PafA species analysed in this study share significant structural homology, *Cgl* PafA diverges significantly at the sequence level, exhibiting only 54.2% identity to its homologs. In contrast, the identities between *Ace*, *Mtb*, and *Msm* PafA exceed 70%. This pronounced sequence divergence likely underpins differences in quaternary structure. Notably, domain-swapping interactions in *Cgl* PafA involve both its N- and C-terminal domains (Pluorde *et al*., 2025), regions that are highly variable compared to the other three species. This suggests *Ace*, *Mtb*, and *Msm* PafA may lack the necessary sequences within their terminal domains to facilitate domain switching, resulting in PafA species that exist stably in a monomeric form. This analysis reinforces a model in which organisms facing persistent environmental stress have evolutionarily favoured a constitutively active, monomeric form of PafA to enable a robust and immediate stress response. In light of this, the dimeric model for *Msm* PafA, derived from mathematical modelling (Ofer et al., 2013), may be subject to reinterpretation, a point that is critically examined in the subsequent section.

4.2 Mg^{2+} Acts as an Allosteric Activator of the Monomeric *Mtb* PafA for enzyme-level regulation

The confirmation of the monomeric state of recombinant *Mtb* PafA prompted an investigation into a potential regulatory mechanism. The exclusive monomeric state of *Mtb* PafA, contrasted with the dormant dimeric form observed in *Cgl* PafA, suggesting that enzyme-level regulation via dimerization does not apply in *Mtb*. This implied the existence of an alternative regulatory mechanism for the PPS in this pathogen; thus an assumption was made that one of the native ligands required for PafA activity may play a vital role in this regulatory mechanism. Based on the findings of Plourde *et al*. (2025), which demonstrated that nucleotide (ATP/ADP) binds and stabilizes the dimeric state of the *Cgl* PafA homologue, which would otherwise slowly dissociate into the monomer state in nucleotide absence (Plourde *et al*., 2025), a compelling hypothesis emerged. The hypothesis posits that Mg^{2+}, a native ligand of PafA, functions as a critical allosteric regulator in *Cgl* PafA, counteracting nucleotide-mediated stabilization to shift the equilibrium from an inactive dimer toward an active monomer. This model was additionally supported by the observation that Pup^{E}—a ligand which requires both ATP and Mg^{2+} for activation—binds to *Msm* PafA independently and non-cooperatively (Ofer *et al*., 2013). Based on this, the present study sought to determine whether Mg^{2+} similarly exerts an allosteric regulatory role in *Mtb* PafA and, if so, whether this regulation influences its oligomeric state under the conditions tested.

Initial SEC analysis of Ni^{2+}-NTA purified PafA revealed that a high concentration of Mg^{2+}—equivalent to that used in optimized activity assay buffers—significantly altered the protein's elution profile. While the majority of PafA population continued to elute at a volume consistent with the monomeric form (54 kDa), distinct protein bands now appeared in the void volume—suggesting the presence of PafA oligomers. Notably, however, this high-molecular-weight void volume fraction contained not only PafA but also the 73 kDa contaminant (presumably the chaperone DnaK), which co-purified with PafA. The co-elution of these proteins in the void volume indicated the formation of a stable PafA complex with a combined mass exceeding the exclusion limit of the resin (>80 kDa), rather than PafA homodimerization. This finding suggested that Mg^{2+} binding induces global conformational changes in monomeric *Mtb* PafA, activating its substrate-binding domain and enabling its access for a specific interaction with DnaK. This interaction holds significant physiological relevance, given that DnaK is a confirmed PPS substrate in both *Msm* and *E. coli* (Watrous *et al*., 2010; Cerda-Maira *et al*., 2011). Thus, the Mg^{2+}-dependent recruitment of DnaK observed likely represents the formation

of a catalytically competent enzyme-substrate complex. This suggests a novel allosteric mechanism for regulating the PPS cascade, potentially controlled by intracellular Mg^{2+} availability. This would represent a form of enzyme-level regulation, a concept with precedent in mycobacteria; for instance, salicylate synthase, essential for siderophore biosynthesis in *Mtb*, is similarly regulated by Mg^{2+} concentration (Mori *et al.*, 2020; Harrison *et al.*, 2006). This Mg^{2+} facilitated regulatory mechanism was directly supported by the control PafA thermal shift assay optimised in this study, revealing key interactions of the protein with ligands in a Mg^{2+} dose-dependent manner. The physiological relevance of such a mechanism is underscored by the critical role of Mg^{2+} in *Mtb* pathogenesis. The bacterium critically requires high intracellular Mg^{2+} concentrations to grow under the mildly acidic conditions (pH 6.0–6.5) it encounters within the macrophage—an environment that is paradoxically Mg^{2+}-deplete (Piddington *et al.*, 2000). Therefore, a model was proposed where the heightened demand for Mg^{2+} during acid stress serves not only as a general nutrient but also as a specific signal to modulate key stress-response pathways. In this context, upon upregulation of Mg^{2+} uptake during primary chronic infection, Mg^{2+} sudden increase would act as an allosteric activator of PafA, activating the PPS to mitigate acid-induced protein damage and/or involvement in DNA damage stress response, both being known triggers for PPS upregulation.

The concept of PafA allostery is not without precedent. As previously mentioned, a foundational model for PafA allostery has been proposed by Ofer and collaborators who concluded—through mathematical modelling—that *Msm* PafA is regulated by PPS protein target substrate binding, which in turn promotes the cooperative binding of Pup (Ofer *et al.*, 2013). A critical re-evaluation of their experimental conditions, however, revealed that Mg^{2+} was a constant, overlooked component of the assay buffer. This suggests that the allosteric behaviour they observed may not be solely attributable to substrate binding but could instead be a consequence of Mg^{2+}-induced conformational changes that prime the enzyme for subsequent substrate and Pup^{E} recruitment. Furthermore, while Ofer *et al*. (2013) mathematically modelled *Msm* PafA as a functional homodimer—a state corroborated by structural studies of the *Cgl* homologue (Özcelik *et al.*, 2012)—biophysical data in this study conclusively demonstrate that *Mtb* PafA is in the monomeric active state (Plourde *et al*., 2025). This discrepancy highlights the potential species-specific differences or condition-dependent oligomerization. While Ofer *et al*. (2013) concluded that *Msm* PafA is a dimer, this assertion merits further scrutiny through methods such as structural visualization. The protein's requirement to recognize varied substrates implies an inherent functional plasticity, supporting

the proteins own self-modification. a property underscored by its self-pupylation activity. Furthermore, the dimeric state may not represent the primary active form, as recent work by Plourde *et al*. (2025) has demonstrated that in *Cgl* PafA, represents dormancy albeit with minimal activity.

The Mg^{2+}-dependent regulation of PafA is physiologically relevant, as both Mg^{2+} acquisition and the PPS are co-upregulated in *M. tuberculosis* during chronic macrophage infection (Hameem and Fatima, 2019). This coordinated response likely represents an essential adaptation for prolonged intracellular survival and the maintenance of virulence under duress. The observed monomeric state of *Mtb* PafA reflects an evolutionary adaptation to the bacterium's obligate pathogenic lifestyle. In this model, the enzyme is maintained in a primed, monomeric state, enabling rapid, single-step activation upon Mg^{2+} rapid acquisition to meet the urgent demands of host infection. This stands in contrast to the reported regulatory mechanism for the non-pathogenic homolog *Cgl* PafA, which exist as dormant dimers that would require a two-step process of dimer dissociation followed by allosteric activation. The streamlined mechanism in *Mtb* potentially confers a fitness advantage by allowing for a more immediate proteotoxic stress response within the hostile host environment.

4.3 PafA is a ligandable target with high structural plasticity

The ligandability of PafA was assessed using a complementary two-stage screening approach. A focused set of seven small-molecule compounds with lead-like properties, sourced from the ChemBridge DIVERSet chemical library, was subjected to both biophysical and cell-based screening.

The thermal stability of recombinant *Mtb* PafA was assessed using a TSA to establish a baseline for its structural integrity. The resulting melt profile indicated a melting temperature (T_m) of 54.08 ± 0.22 °C, a value marginally below the estimated range of 55 - 60 °C for wild-type *Mtb* PafA and lower than the ~57 °C reported by Li *et al*. (2025) for a mutant variant of *Mtb* PafA (Li *et al.*, 2025). This finding aligned well with the stated objective of the Li *et al*. (2025) mutations, which was to enhance the solubility of recombinant *Mtb* PafA. While Li and coworkers introduced point mutations to facilitate soluble expression of *Mtb* PafA, a goal reflected in the modest T_m increase of their variant, the present study achieved high yields of soluble, wild-type PafA through chaperone-assisted expression. This optimization emphasised

that robust recombinant production is attainable without genetic manipulation, thereby preserving the protein's native sequence and, presumably, its authentic regulatory and catalytic properties.

Importantly, the thermal denaturation profile of *Mtb* PafA provided strong biophysical evidence consistent with a monomeric protein state. The assay yielded a classic, monophasic unfolding curve indicative of a two-state transition between a folded native state and an unfolded denatured state. This is in direct contrast to the biphasic melt profiles characteristic of typical oligomeric proteins, particularly dimers, which undergo a three-state unfolding process involving dissociation into monomers followed by the unfolding of the monomeric units themselves (Burgos *et al.*, 2008). The absence of a second transition phase strongly corroborated the protein's biophysical data that the native, folded state of recombinant *Mtb* PafA is a monomer. To validate the optimized thermal TSA for ligand-binding studies, PafA interaction with its native substrate ATP was characterized. This ligand served as an ideal positive control, as its binding is integral to the enzyme's catalytic mechanism. Pupylation proceeds via a two-step reaction catalyzed by PafA: first, the binding and phosphorylation of the C-terminal glutamate of Pup^E, and second, the subsequent ligation of the activated Pup^E to a lysine residue on the target protein (Guth *et al.*, 2011). This mechanistic framework supports a model wherein PafA is primed for catalysis through the cooperative binding of Mg^{2+} and ATP (Mg^{2+}-ATP complex), which induces an allosteric transition from a basal state to an activated conformation. This activated state is in turn competent for binding and phosphorylating the C-terminal glutamate of Pup^E. Consequently, the binding of ATP—a cornerstone of this activation mechanism— was predicted to confer significant conformational stability to PafA. This established the theoretical basis for using ATP-induced thermal stabilization as a benchmark to validate the assay and to assess the binding of novel ligands.

Integrating these findings lead to a model in which Mg^{2+}, rather than ATP, as the principal allosteric activator of *Mtb* PafA. While Plourde *et al*. (2025) demonstrated that ATP binding stabilizes the dormant dimeric state of the *Cgl* homologue, Mg^{2+} binding to the monomeric *Mtb* PafA induced a conformational change that facilitated substrate engagement. This regulatory mechanism is postulated to also be conserved in the *Clg* and *Msm* PafA homologs, as prior independent evaluations of their allosteric properties overlooked the presence of Mg^{2+} in the respective buffer compositions. Therefore, a model was proposed wherein Mg^{2+} acts as the allosteric activator of PafA, priming the enzyme for efficient ATP binding. The resulting Mg^{2+}-

ATP-PafA complex subsequently activates the Pup^E substrate for transfer. This model was tested by evaluating the ability of Mg^{2+} to enhance the affinity and stability of the PafA-ATP interaction. Analysis of the thermal shift data revealed that neither ATP nor Mg^{2+} alone induced a significant stabilizing effect on apo *Mtb* PafA ($p > 0.05$), indicating weak independent binding. This was consistent with reports that ATP binds to the dormant dimeric state of *Cgl* PafA, a low-affinity interaction that the presented data suggests is conserved in monomeric *Mtb* PafA (Plourde *et al.*, 2025). Notably, however, pre-incubation of PafA with Mg^{2+} prior to ATP addition resulted in a pronounced, concentration-dependent stabilization, with T_m increases of up to 2.22 °C. This strong synergistic stabilization suggested that Mg^{2+} binding induces a conformational change in PafA that dramatically increases its affinity for ATP, forming Mg^{2+}-ATP complexes that effectively remodel the active site into a high-affinity state primed for catalysis.

This model of Mg^{2+}-facilitated ATP binding was further supported by in-silico and structural evidence. Molecular docking of ATP to monomeric *Cgl* PafA yielded a more favourable score in the presence of Mg^{2+} (-11.278) than in its absence (-9.909), indicating a stronger predicted binding affinity that aligns with the observed thermal stabilization. Furthermore, the first resolved crystal structure of *Cgl* PafA (4B0T), was solved as a homodimer with ADP and Mg^{2+} bound in the active site of each monomer. This suggested that the enzyme possesses an intrinsic affinity for these ligands and that strong binding of ATP to the active site requires Mg^{2+} presence, reinforcing the physiological relevance of the Mg^{2+}-nucleotide complex. (Özcelik *et al.*, 2012). Collectively, these findings support a model in which Mg^{2+} acts as a master allosteric regulator of *Mtb* PafA. Mg^{2+} binding induces a conformational shift that activates the enzyme through two coordinated mechanisms: first, by exposing the target protein binding site, as evidenced by the Mg^{2+}-dependent recruitment of DnaK; and second, by restructuring the active site to favour ATP binding over ADP. This switch in nucleotide preference is critical for initiating the pupylation reaction. The thermal shift assays provide direct biophysical support for this model: while ATP alone had a negligible effect on apo *Mtb* PafA stability ($\Delta T_m = +0.26$ °C, $p > 0.05$), the addition of Mg^{2+} triggered a pronounced stabilization, consistent with high-affinity ATP binding following the displacement of endogenous ADP. This Mg^{2+}-dependent enhancement of ATP affinity is further validated by in-silico analysis, which calculated a more favourable binding free energy (MM/GBSA ΔGbind) for ATP (-47.51 kJ/mol) than for ADP (-35.90 kJ/mol) in the presence of the cofactor.

Using the optimized TSA protocol, all seven candidate ligands significantly stabilized *Mtb* PafA in solution at the optimal Mg^{2+} concentration (Compound C11 was excluded from this screen due to suspected promiscuous binding identified in a prior integrity assay; see Appendix Figure S4). This stabilization was corroborated by the CETSA, which revealed a congruent stabilising effect against thermal denaturation for the same ligands. In CETSA, these compounds demonstrated high affinity, increasing the protein's thermal stability by more than 8 °C relative to the untreated control. The inherent substrate plasticity of PafA, which enables it to target a diverse "pupylome," may account for the identification of this relatively large set of stabilizing ligands. Among the screened ligands, C10 emerged as the most potent *Mtb* PafA binder. It conferred the greatest thermal stabilization in the cellular CETSA assay (ΔT_m = +10.29 °C) and was one of the top two stabilizers in the biophysical TSA (ΔT_m = +1.10 °C), an effect comparable to that of the native ATP ligand at high millimolar concentration (ΔT_m = +1.33 °C). Crucially, like ATP, all ligands required Mg^{2+} to bind PafA significantly *in vitro*, suggesting a shared binding mechanism. This was confirmed by computational modelling, which showed all compounds preferentially occupying the ATP-binding pocket. Analysis of the docked poses correlated with the *in vitro* efficacy, identifying A3, C10, and D2 as the top-ranked binders—consistent with their top performance in the thermal stability assays. The models further indicated that Mg^{2+} coordination is essential for stabilizing a rigid ligand-protein complex, in contrast to the increased conformational flexibility observed in its absence.

The preferential binding of the compounds to the ATP-binding site initially suggested they might act as PafA inhibitors, potentially blocking ATP access and thereby inhibiting the PPS pathway. To determine the functional impact of the compounds, mycobacterial proliferation was assessed under both standard culture conditions and in the presence of induced nitrosative stress. Intriguingly, compounds A3, C9, and D3 significantly increased mycobacterial cell growth under standard culture conditions, indicating a stimulatory effect. This indicated that the compounds may bind and stabilize endogenous *Msm* PafA, thereby enhancing its basal activity. Consequently, a model is proposed in which the ligands bind to the ATP-occupied form of PafA, cooperatively stabilizing the complex and amplifying its housekeeping function. This model is supported by the CETSA results, which showed that all the candidate PafA ligands bound recombinant PafA even within the cellular milieu, which is rich in ATP and Mg^{2+}. These findings further support the hypothesis that the resolved crystal structures capture a dormant conformation. It is plausible that Mg^{2+} induces a global conformational change in

PafA, significantly altering the binding landscape to favour the cooperative binding of both ATP and the identified ligands.

The lack of growth inhibition by the candidate compounds under standard conditions was anticipated, indicating specificity for the PPS pathway without pleiotropic toxicity. This contrasted with the significant inhibition caused by the control compound ST1926, which aligns with its reported anti-mycobacterial activity (Li *et al*., 2022). Since the PPS is primarily upregulated during stress, its inhibition is not expected to affect growth under optimal conditions (Guzzo *et al*., 2021; von Rosen *et al*., 2021). Therefore, the observed growth inhibition by ST1926 under standard conditions likely reflects off-target effects on essential cellular processes. Under nitrosative stress (2.5 mM $NaNO_2$), the candidate compounds had no significant effect on proliferation relative to the stressed control. This suggested that they do not function as canonical PafA inhibitors under these conditions. However, it is possible that the chosen stressor concentration was overly potent, causing severe growth suppression that could mask any additional inhibitory effect from PPS pathway modulation

4.3 Summary and future direction

To conclude, this study successfully achieved the soluble overexpression and purification of wild-type *Mtb* PafA in its native conformation without the use of denaturants or solubility-enhancing mutations. Biophysical characterization of *Mtb* PafA revealed it to be a stable monomer whose activity is allosterically regulated by the protein's metal cofactor Mg^{2+}. Furthermore, seven novel ligands conferring significant thermal stabilization upon binding were identified for *Mtb* PafA and its homologs, which showed no inhibitory activity under the tested conditions. These findings indicate that all identified ligands function as PafA stabilizers, revealing their potential as candidates for novel BacPROTAC development within the PPS. Notably, ligand C10, A2, D2, and A3 emerged as the most promising candidates, conferring the greatest cross-assay activity (TSA/CETSA/docking) to the protein and representing a prime scaffold for future optimization (Table 4.1).

Table 4.1. Cross assay binding activity of PafA ligands

Compound	ΔT_m (°C)	ΔTagg (°C)	Docking score
ATP	**2.22**	**-**	**-11.278**
A2	**1.20**	**8.60**	**-6.44**
A3	**0.85**	**9.31**	**-8.338**
B7	0.90	7.46	-4.876
C9	0.90	7.85	-5.286
C10	**1.10**	10.29	**-8.474**
D2	0.75	**8.48**	**-8.592**
D3	0.90	8.47	-4.791

While this study presents key findings, critical limitations warrant consideration. A principal constraint was the reliance on the dimeric *Cgl* PafA structures (PDB: 4B0T, 4BJR) for all *in silico* analyses. Although these homologs share significant structural similarity with *Mtb* PafA, they likely represent a low-affinity, dormant conformation. The Mg^{2+}-dependent priming mechanism proposed here suggests that cofactor binding induces global conformational changes, generating a distinct, high-affinity active state. This discrepancy may explain the non-inhibitory nature of the ligands: rather than competitively displacing ATP, they may bind cooperatively to a reorganized site within the Mg^{2+}-primed enzyme, as suggested by CETSA data showing binding in an ATP-replete cellular environment. Consequently, a critical future direction is the structural determination of monomeric, Mg^{2+}-bound PafA, particularly from *Mtb*. Such structures would provide an accurate template for computational modelling of ligand interactions within the physiologically relevant, primed active site, thereby resolving the mechanistic paradox between binding location and functional outcome. Furthermore, future work should transition to endogenous PafA in mycobacterial hosts, specifically *Msm* and, crucially, *Mtb*. This would directly assess and validate compound permeability and their binding to native, untagged protein within its physiological cellular environment, eliminating any potential confounding influence of the recombinant His-tag—despite prior evidence suggesting it does not impair activity (Cerda-Maira *et al*., 2011).

REFERENCES

Ahmad, M., Aduru, S.V., Smith, R.P., Zhao, Z. and Lopatkin, A.J., 2025. The role of bacterial metabolism in antimicrobial resistance. *Nature Reviews Microbiology*, *23*(7), pp.439-454.

Ahmad, S., 2011. Pathogenesis, immunology, and diagnosis of latent Mycobacterium tuberculosis infection. *Journal of Immunology Research*, *2011*(1), pp.814943.

Ahn, D.I. and World Health Organization, 2005. Addressing poverty in TB control: options for national TB control programmes. In *Addressing Poverty in TB Control: Options For National TB Control Programmes*, pp.80.

Barandun, J., Delley, C.L. and Weber-Ban, E., 2012. The Pupylation Pathway and Its Role in mycobacteria. *BMC Biology*, *10*(1), pp.95.

Barandun, J., Delley, C.L., Ban, N. and Weber-Ban, E., 2013. Crystal structure of the complex between prokaryotic ubiquitin-like protein and its ligase PafA. *Journal of the American Chemical Society*, *135*(18), pp.6794-6797.

Bond, M.J. and Crews, C.M., 2021. Proteolysis targeting chimeras (PROTACs) come of age: entering the third decade of targeted protein degradation. *RSC Chemical Biology*, *2*(3), pp.725-742.

Bondeson, D.P. and Crews, C.M., 2017. Targeted protein degradation by small molecules. *Annual Review of Pharmacology and Toxicology*, *57*(1), pp.107-123.

Bondeson, D.P., Smith, B.E., Burslem, G.M., Buhimschi, A.D., Hines, J., Jaime-Figueroa, S., Wang, J., Hamman, B.D., Ishchenko, A. and Crews, C.M., 2018. Lessons in PROTAC design from selective degradation with a promiscuous warhead. *Cell Chemical Biology*, *25*(1), pp.78-87.

Brites, D. and Gagneux, S., 2017. The nature and evolution of genomic diversity in the Mycobacterium tuberculosis complex. *Strain variation in the Mycobacterium tuberculosis complex: its role in biology, epidemiology and control*, pp.1-26.

Bryant, J.M., Brown, K.P., Burbaud, S., Everall, I., Belardinelli, J.M., Rodriguez-Rincon, D., Grogono, D.M., Peterson, C.M., Verma, D., Evans, I.E. and Ruis, C., 2021. Stepwise pathogenic evolution of *Mycobacterium* abscessus. *Science*, *372*(6541), p.eabb8699.

Burgos, I., Dassie, S.A. and Fidelio, G.D., 2008. Thermodynamic model for the analysis of calorimetric data of oligomeric proteins. *The Journal of Physical Chemistry B*, *112*(45), pp.14325-14333.

Cerda-Maira, F.A., McAllister, F., Bode, N.J., Burns, K.E., Gygi, S.P. and Darwin, K.H., 2011. Reconstitution of the *Mycobacterium* tuberculosis pupylation pathway in Escherichia coli. *EMBO Reports*, *12*(8), pp.863-870.

Chen, X., Qiu, J.D., Shi, S.P., Suo, S.B. and Liang, R.P., 2013. Systematic analysis and prediction of pupylation sites in prokaryotic proteins. *PloS One*, *8*(9), pp.e74002.

Chetty, R., Delport, A., Mthembu, S., Veale, C.G. and Hewer, R., 2025. Screening the Pandemic Response Box identifies novel ligands of the Staphylococcus aureus protein arginine kinase, McsB. *Molecular Biology Reports*, *52*(1), pp.446.

Cui, H., Müller, A.U., Leibundgut, M., Tian, J., Ban, N. and Weber-Ban, E., 2021. Structures of prokaryotic ubiquitin-like protein Pup in complex with depupylase Dop reveal the mechanism of catalytic phosphate formation. *Nature Communications*, *12*(1), pp.6635.

Danchik, C., Wang, S. and Karakousis, P.C., 2021. Targeting the mycobacterium tuberculosis stringent response as a strategy for shortening tuberculosis treatment. *Frontiers in microbiology*, *12*, pp.744167.

Darwin, K.H., Lin, G., Chen, Z., Li, H. and Nathan, C.F., 2005. Characterization of a *Mycobacterium* tuberculosis proteasomal ATPase homologue. *Molecular Microbiology*, *55*(2), pp.561-571.

Delley, C.L., Striebel, F., Heydenreich, F.M., Özcelik, D. and Weber-Ban, E., 2012. Activity of the mycobacterial proteasomal ATPase Mpa is reversibly regulated by pupylation. *Journal of Biological Chemistry*, *287*(11), pp.7907-7914.

Delport, A. and Hewer, R., 2022. A superior loading control for the cellular thermal shift assay. *Scientific Reports*, *12*(1), pp.6672.

Diricks, M., Maurer, F.P., Dreyer, V., Barilar, I., Utpatel, C., Merker, M., Wetzstein, N. and Niemann, S., 2025. Genomic insights into the plasmidome of non-tuberculous mycobacteria. *Genome Medicine*, *17*(1), pp.19.

Dorman, S.E., Nahid, P., Kurbatova, E.V., Goldberg, S.V., Bozeman, L., Burman, W.J., Chang, K.C., Chen, M., Cotton, M., Dooley, K.E. and Engle, M., 2020. High-dose rifapentine with or without moxifloxacin for shortening treatment of pulmonary tuberculosis: study protocol for TBTC study 31/ACTG A5349 phase 3 clinical trial. *Contemporary Clinical Trials*, *90*, pp.105938.

Elharar, Y., Roth, Z., Hecht, N., Rotkopf, R., Khalaila, I. and Gur, E., 2016. Posttranslational regulation of coordinated enzyme activities in the Pup-proteasome system. *Proceedings of the National Academy of Sciences*, *113*(12), pp.E1605-E1614.

Festa, R.A., McAllister, F., Pearce, M.J., Mintseris, J., Burns, K.E., Gygi, S.P. and Darwin, K.H., 2010. Prokaryotic ubiquitin-like protein (Pup) proteome of Mycobacterium tuberculosis. *PloS One*, *5*(1), pp.e8589.

Festa, R.A., Pearce, M.J. and Darwin, K.H., 2007. Characterization of the proteasome accessory factor (paf) operon in Mycobacterium tuberculosis. *Journal of Bacteriology*, *189*(8), pp.3044-3050.

Fleming, J., Magana, P., Nair, S., Tsenkov, M., Bertoni, D., Pidruchna, I., Afonso, M.Q.L., Midlik, A., Paramval, U., Žídek, A. and Laydon, A., 2025. AlphaFold Protein Structure Database and 3D-Beacons: New Data and Capabilities. *Journal of Molecular Biology*, p.168967.

Fryland, T., Christensen, J.H., Pallesen, J., Mattheisen, M., Palmfeldt, J., Bak, M., Grove, J., Demontis, D., Blechingberg, J., Ooi, H.S. and Nyegaard, M., 2016. Identification of the BRD1 interaction network and its impact on mental disorder risk. *Genome Medicine*, *8*(1), p.53.

Fudrini Olivencia, B., Müller, A.U., Roschitzki, B., Burger, S., Weber-Ban, E. and Imkamp, F., 2017. Mycobacterium smegmatis PafBC is involved in regulation of DNA damage response. *Scientific Reports*, *7*(1), p.13987.

Gagneux, S., 2018. Ecology and evolution of Mycobacterium tuberculosis. *Nature Reviews Microbiology*, *16*(4), pp.202-213.

Gandotra, S., Schnappinger, D., Monteleone, M., Hillen, W. and Ehrt, S., 2007. *In vivo* gene silencing identifies the Mycobacterium tuberculosis proteasome as essential for the bacteria to persist in mice. *Nature Medicine*, *13*(12), pp.1515-1520.

Goletti, D., Meintjes, G., Andrade, B.B., Zumla, A. and Lee, S.S., 2025. Insights from the 2024 WHO global tuberculosis report–more comprehensive action, innovation, and investments required for achieving WHO end TB goals. *International Journal of Infectious Diseases*, *150*, pp.107325

Gulumbe, B.H., Abdulrahim, A., Ahmad, S.K., Lawan, K.A. and Danlami, M.B., 2025. WHO report signals tuberculosis resurgence: Addressing systemic failures and revamping control strategies. *Decoding Infection and Transmission*, *3*, pp.100044.

Guth, E., Thommen, M. and Weber-Ban, E., 2011. Mycobacterial ubiquitin-like protein ligase PafA follows a two-step reaction pathway with a phosphorylated pup intermediate. *Journal of Biological Chemistry*, *286*(6), pp.4412-4419.

Guzzo, M.B., Li, Q., Nguyen, H.V., Boom, W.H. and Nguyen, L., 2021. The pup-proteasome system protects mycobacteria from antimicrobial antifolates. *Antimicrobial Agents and Chemotherapy*, *65*(4), pp.10-1128.

Hameed, S. and Fatima, Z., 2019. Targeting Magnesium Homeostasis as Potential Anti-infective Strategy Against mycobacteria. In *Frontiers in Clinical Drug Research-Anti Infectives, 5,* pp.212-223.

Harrison, A.J., Yu, M., Gårdenborg, T., Middleditch, M., Ramsay, R.J., Baker, E.N. and Lott, J.S., 2006. The structure of MbtI from Mycobacterium tuberculosis, the first enzyme in the biosynthesis of the siderophore mycobactin, reveals it to be a salicylate synthase. *Journal of Bacteriology*, *188*(17), pp.6081-6091.

Iyer, L.M., Burroughs, A.M. and Aravind, L., 2008. Unraveling the biochemistry and provenance of pupylation: a prokaryotic analog of ubiquitination. *Biology Direct*, *3*(1), p.45.

Izert-Nowakowska, M.A., Klimecka, M.M., Antosiewicz, A., Wróblewski, K., Kowalski, J.J., Bandyra, K.J., Góral, T., Kmiecik, S., Serwa, R.A. and Górna, M.W., 2025. Targeted protein degradation in Escherichia coli using CLIPPERs. *EMBO Reports*, pp.1-23.

Janssens, J.P. and Rieder, H.L., 2008. An ecological analysis of incidence of tuberculosis and per capita gross domestic product. *European Respiratory Journal*, *32*(5), pp.1415-1416.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A. and Bridgland, A., 2021. Highly accurate protein structure prediction with AlphaFold. *Nature*, *596*(7873), pp.583-589.

Kestler, B. and Tyler, S.K., 2022. Latent tuberculosis testing through the ages: the search for a sleeping killer. *American Journal of Physiology-Lung Cellular and Molecular Physiology*, *322*(3), pp.L412-L419.

Klein, K., Kato, M., Frank-Bertoncelj, M., Kolling, C., Ciurea, A., Gay, S. and Ospelt, C., 2018. Evaluating the bromodomain protein BRD1 as a therapeutic target in rheumatoid arthritis. *Scientific Reports*, *8*(1), p.11125.

Koch, C.P., Perna, A.M., Weissmüller, S., Bauer, S., Pillong, M., Baleeiro, R.B., Reutlinger, M., Folkers, G., Walden, P., Wrede, P. and Hiss, J.A., 2013. Exhaustive proteome mining for functional MHC-I ligands. *ACS Chemical Biology*, *8*(9), pp.1876-1881.

Kock, R., Michel, A.L., Yeboah-Manu, D., Azhar, E.I., Torrelles, J.B., Cadmus, S.I., Brunton, L., Chakaya, J.M., Marais, B., Mboera, L. and Rahim, Z., 2021. Zoonotic tuberculosis–the changing landscape. *International Journal of Infectious Diseases*, *113*, pp.S68-S72.

Lamichhane, G., Raghunand, T.R., Morrison, N.E., Woolwine, S.C., Tyagi, S., Kandavelou, K. and Bishai, W.R., 2006. Deletion of a Mycobacterium tuberculosis proteasomal ATPase homologue gene produces a slow-growing strain that persists in host tissues. *The Journal of Infectious Diseases*, *194*(9), pp.1233-1240.

Lavania, M., Singh, I., Turankar, R.P., Ahuja, M., Pathak, V., Sengupta, U., Das, L., Kumar, A., Darlong, J., Nathan, R. and Maseey, A., 2018. Molecular detection of multidrug-

resistant Mycobacterium leprae from Indian leprosy patients. *Journal of Global Antimicrobial Resistance, 12, pp.* 214-219.

Li, C., Liu, S., Dong, B., Li, C., Jian, L., He, J., Zeng, J., Zhou, Q., Jia, D., Luo, Y. and Sun, Q., 2022. Discovery and mechanistic study of *Mycobacterium tuberculosis* PafA inhibitors. *Journal of Medicinal Chemistry*, *65*(16), pp.11058-11065.

Li, W., Matsuoka, M., Kai, M., Thapa, P., Khadge, S., Hagge, D.A., Brennan, P.J. and Vissa, V., 2012. Real-time PCR and high-resolution melt analysis for rapid detection of Mycobacterium leprae drug resistance mutations and strain types. *Journal of Clinical Microbiology*, *50*(3), pp.742-753.

Li, X., Li, G., Yang, J., Jin, G., Shao, Y., Li, Y., Wei, P. and Zhang, L., 2022. Drug resistance (dapsone, rifampicin, ofloxacin) and resistance-related gene mutation features in leprosy patients: a systematic review and meta-analysis. *International Journal of Molecular Sciences*, *23*(20), p.12443.

Liu, J., Ma, J., Liu, Y., Xia, J., Li, Y., Wang, Z.P. and Wei, W., 2020, December. PROTACs: a novel strategy for cancer therapy. In *Seminars in Cancer Biology*, 67, pp.171-179.

Mabanglo, M.F. and Houry, W.A., 2022. Recent structural insights into the mechanism of ClpP protease regulation by AAA+ chaperones and small molecules. *Journal of Biological Chemistry*, *298*(5), pp.101781.

Melkani, G.C., Zardeneta, G. and Mendoza, J.A., 2003. The ATPase activity of GroEL is supported at high temperatures by divalent cations that stabilize its structure. *Biometals*, *16*(3), pp.479-484.

Mohammadnabi, N., Shamseddin, J., Emadi, M., Bodaghi, A.B., Varseh, M., Shariati, A., Rezaei, M., Dastranj, M. and Farahani, A., 2024. Mycobacterium tuberculosis: the mechanism of pathogenicity, immune responses, and diagnostic challenges. *Journal of clinical laboratory analysis*, *38*(23), pp.e25122.

Mori, M., Stelitano, G., Gelain, A., Pini, E., Chiarelli, L.R., Sammartino, J.C., Poli, G., Tuccinardi, T., Beretta, G., Porta, A. and Bellinzoni, M., 2020. Shedding X-ray light on the

role of magnesium in the activity of Mycobacterium tuberculosis salicylate synthase (MbtI) for drug design. *Journal of Medicinal Chemistry*, *63*(13), pp.7066-7080.

Morreale, F.E., Kleine, S., Leodolter, J., Junker, S., Hoi, D.M., Ovchinnikov, S., Okun, A., Kley, J., Kurzbauer, R., Junk, L. and Guha, S., 2022. BacPROTACs mediate targeted protein degradation in bacteria. *Cell*, *185*(13), pp.2338-2353.

Murdoch, J., Curran, R., van Rensburg, A.J., Awotiwon, A., Dube, A., Bachmann, M., Petersen, I. and Fairall, L., 2021. Identifying contextual determinants of problems in tuberculosis care provision in South Africa: a theory-generating case study. *Infectious Diseases of Poverty*, *10*(03), pp.82-94.

Nahid, P., Dorman, S.E., Alipanah, N., Barry, P.M., Brozek, J.L., Cattamanchi, A., Chaisson, L.H., Chaisson, R.E., Daley, C.L., Grzemska, M. and Higashi, J.M., 2016. Official American thoracic society/centers for disease control and prevention/infectious diseases society of America clinical practice guidelines: treatment of drug-susceptible tuberculosis. *Clinical infectious diseases*, *63*(7), pp.e147-e195.

Nguyen, K.T., Piastro, K., Gray, T.A. and Derbyshire, K.M., 2010. Mycobacterial biofilms facilitate horizontal DNA transfer between strains of Mycobacterium smegmatis. *Journal of Bacteriology*, *192*(19), pp.5134-5142.

Niemann, S. and Supply, P., 2014. Diversity and evolution of Mycobacterium tuberculosis: moving to whole-genome-based approaches. *Cold Spring Harbor perspectives in medicine*, *4*(12), pp.a021188.

Ofer, N., Forer, N., Korman, M., Vishkautzan, M., Khalaila, I. and Gur, E., 2013. Allosteric transitions direct protein tagging by PafA, the prokaryotic ubiquitin-like protein (Pup) ligase. *Journal of Biological Chemistry*, *288*(16), pp.11287-11293.

Oxlade, O. and Murray, M., 2012. Tuberculosis and poverty: why are the poor at greater risk in India? *PloS One*, *7*(11), pp.e47533.

Özcelik, D., Barandun, J., Schmitz, N., Sutter, M., Guth, E., Damberger, F.F., Allain, F.H.T., Ban, N. and Weber-Ban, E., 2012. Structures of Pup ligase PafA and depupylase Dop from the prokaryotic ubiquitin-like modification pathway. *Nature Communications*, *3*(1), pp.1014.

Perera, G., Thilakarathna, M., Aluthgamage, I., Sathsarani, S., Fernando, P.C., Samaranayake, S., Ullah, N. and Deepachandi, B., 2025. Unraveling Drug Resistance in Mycobacterium leprae: Exploring Genetic Mutations to Enhance Treatment Strategies for Human Leprosy—A Narrative Review. *International Journal of Microbiology*, *2025*(1), pp.7204337.

Pettersson, M. and Crews, C.M., 2019. PROteolysis TArgeting Chimeras (PROTACs)—past, present and future. *Drug Discovery Today: Technologies*, *31*, pp.15-27.

Pettit, A.C., Phillips, P.P., Kurbatova, E., Vernon, A., Nahid, P., Dawson, R., Dooley, K.E., Sanne, I., Waja, Z., Mohapi, L. and Podany, A.T., 2023. Rifapentine with and without moxifloxacin for pulmonary tuberculosis in people with human immunodeficiency virus (S31/A5349). *Clinical Infectious Diseases*, *76*(3), pp.e580-e589.

Pfaff, P., Samarasinghe, K.T., Crews, C.M. and Carreira, E.M., 2019. Reversible spatiotemporal control of induced protein degradation by bistable PhotoPROTACs. *ACS Central Science*, *5*(10), pp.1682-1690.

Piddington, D.L., Kashkouli, A. and Buchmeier, N.A., 2000. Growth of Mycobacterium tuberculosis in a defined medium is very restricted by acid pH and Mg2+ levels. *Infection and Immunity*, *68*(8), pp.4518-4522.

Plourde, A., Ogata-Bean, J.C. and Vahidi, S., 2025. Mapping the structural heterogeneity of Pup ligase PafA using H/D exchange mass spectrometry. *Journal of Biological Chemistry*, *301*(5), pp.108437

Poulsen, C., Akhter, Y., Jeon, A.H.W., Schmitt-Ulms, G., Meyer, H.E., Stefanski, A., Stühler, K., Wilmanns, M. and Song, Y.H., 2010. Proteome-wide identification of mycobacterial pupylation targets. *Molecular Systems Biology*, *6*(1), pp.386.

Price, C.T., Hanford, H.E., Al-Quadan, T., Santic, M., Shin, C.J., Da'as, M.S. and Abu Kwaik, Y., 2024. Amoebae as training grounds for microbial pathogens. *MBio*, *15*(8), pp.e00827-24.

Queval, C.J., Brosch, R. and Simeone, R., 2017. The macrophage: a disputed fortress in the battle against Mycobacterium tuberculosis. *Frontiers in microbiology*, *8*, pp.2284.

Regev, O., Linder, H. and Gur, E., 2019. Pup-Click—A New Chemoenzymatic Method for the Generation of Singly Pupylated Targets. *Bioconjugate Chemistry*, *30*(11), pp.2909-2916.

Reis-Santos, B., Shete, P., Bertolde, A., Sales, C.M., Sanchez, M.N., Arakaki-Sanchez, D., Andrade, K.B., Gomes, M.G.M., Boccia, D., Lienhardt, C. and Maciel, E.L., 2019. Tuberculosis in Brazil and cash transfer programs: a longitudinal database study of the effect of cash transfer on cure rates. *PloS One*, *14*(2), pp.e0212617.

Russell, D.W. and Sambrook, J., 2001. *Molecular Cloning: A Laboratory Manual* (Vol. 1, p. 112). Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Sakamoto, K.M., Kim, K.B., Kumagai, A., Mercurio, F., Crews, C.M. and Deshaies, R.J., 2001. Protacs: Chimeric molecules that target proteins to the Skp1–Cullin–F box complex for ubiquitination and degradation. *Proceedings of the National Academy of Sciences*, *98*(15), pp.8554-8559.

Schilling, C.M., Zdanowicz, R., Rabl, J., Müller, A.U., Boehringer, D., Glockshuber, R. and Weber-Ban, E., 2025. Single-stranded DNA binding to the transcription factor PafBC triggers the mycobacterial DNA damage response. *Science Advances*, *11*(6), p.eadq9054.

Schneekloth, A.R., Pucheault, M., Tae, H.S. and Crews, C.M., 2008. Targeted intracellular protein degradation induced by a small molecule: En route to chemical proteomics. *Bioorganic & Medicinal Chemistry Letters*, *18*(22), pp.5904-5908.

Sekar, B., Arunagiri, K., Narayanan, S., Menaka, K. and Oommen, P.K., 2011. Detection of mutations in folp1, rpoB and gyrA genes of M. leprae by PCR-direct sequencing–a rapid tool for screening drug resistance in leprosy. *Leprosy review*, *82*(1), pp.36-45.

Shah, N.S., Wright, A. and Drobniewski, F., 2005. Extreme drug resistance in tuberculosis (XDR-TB): global survey of supranational reference laboratories for Mycobacterium tuberculosis with resistance to second-line drugs. *Int J Tuberc Lung Dis*, *9*(1), pp.S77.

Striebel, F., Hunkeler, M., Summer, H. and Weber-Ban, E., 2010. The mycobacterial Mpa–proteasome unfolds and degrades pupylated substrates by engaging Pup's N-terminus. *The EMBO Journal*, *29*(7), pp.1262-1271.

Striebel, F., Imkamp, F., Özcelik, D. and Weber-Ban, E., 2014. Pupylation as a signal for proteasomal degradation in bacteria. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, *1843*(1), pp.103-113.

Sutter, M., Damberger, F.F., Imkamp, F., Allain, F.H.T. and Weber-Ban, E., 2010. Prokaryotic ubiquitin-like protein (Pup) is coupled to substrates via the side chain of its C-terminal glutamate. *Journal of the American Chemical Society*, *132*(16), pp.5610-5612.

Telenti, A., Imboden, P., Marchesi, F., Matter, L., Schopfer, K., Bodmer, T., Lowrie, D., Colston, M.J. and Cole, S., 1993. Detection of rifampicin-resistance mutations in Mycobacterium tuberculosis. *The Lancet*, *341*(8846), pp.647-651.

Ummels, R., Abdallah, A.M., Kuiper, V., Aâjoud, A., Sparrius, M., Naeem, R., Spaink, H.P., van Soolingen, D., Pain, A. and Bitter, W., 2014. Identification of a novel conjugative plasmid in mycobacteria that requires both type IV and type VII secretion. *MBio*, *5*(5), pp.10-1128.

Van Der Laan, R., Snabilié, A. and Obradovic, M., 2022. Meeting the challenges of NTM-PD from the perspective of the organism and the disease process: innovations in drug development and delivery. *Respiratory Research*, *23*(1), pp.376.

Velayati, A.A., Masjedi, M.R., Farnia, P., Tabarsi, P., Ghanavi, J., ZiaZarifi, A.H. and Hoffner, S.E., 2009. Emergence of new forms of totally drug-resistant tuberculosis bacilli: super extensively drug-resistant tuberculosis or totally drug-resistant strains in Iran. *Chest*, *136*(2), pp.420-425.

Veyrier, F.J., Dufort, A. and Behr, M.A., 2011. The rise and fall of the Mycobacterium tuberculosis genome. *Trends in Microbiology*, *19*(4), pp.156-161.

Von Rosen, T., Keller, L.M. and Weber-Ban, E., 2021. Survival in hostile conditions: pupylation and the proteasome in actinobacterial stress response pathways. *Frontiers in Molecular Biosciences*, *8*, pp.685757.

Watrous, J., Burns, K., Liu, W.T., Patel, A., Hook, V., Bafna, V., Barry 3rd, C.E., Bark, S. and Dorrestein, P.C., 2010. Expansion of the mycobacterial "PUPylome". *Molecular BioSystems*, *6*(2), pp.376-385.

Winter, G.E., Buckley, D.L., Paulk, J., Roberts, J.M., Souza, A., Dhe-Paganon, S. and Bradner, J.E., 2015. Phthalimide conjugation as a strategy for *in vivo* target protein degradation. *Science*, *348*(6241), pp.1376-1381.

Xia, X., 2023. Horizontal Gene transfer and drug resistance involving Mycobacterium tuberculosis. *Antibiotics*, *12*(9), pp.1367.

Yimcharoen, M., Saikaew, S., Wattananandkul, U., Phunpae, P., Intorasoot, S., Tayapiwatana, C. and Butr-Indr, B., 2023. Mycobacterium tuberculosis adaptation in response to isoniazid treatment in a multi-stress system that mimics the host environment. *Antibiotics*, *12*(5), pp.852.

Zainuddin, Z.F. and Dale, J.W., 1990. Does Mycobacterium tuberculosis have plasmids? *Tubercle*, *71*(1), pp.43-49.

Zhang, S., Burns-Huang, K.E., Janssen, G.V., Li, H., Ovaa, H., Hedstrom, L. and Darwin, K.H., 2017. Mycobacterium tuberculosis proteasome accessory factor A (PafA) can transfer prokaryotic ubiquitin-like protein (pup) between substrates. *MBio*, *8*(1), pp.10-1128.

Zwyer, M., Rutaihwa, L.K., Windels, E., Hella, J., Menardo, F., Sasamalo, M., Sommer, G., Schmülling, L., Borrell, S., Reinhard, M. and Dötsch, A., 2023. Back-to-Africa introductions of Mycobacterium tuberculosis as the main cause of tuberculosis in Dar es Salaam, Tanzania. *PLoS Pathogens*, *19*(4), pp.1010893.

APPENDIX

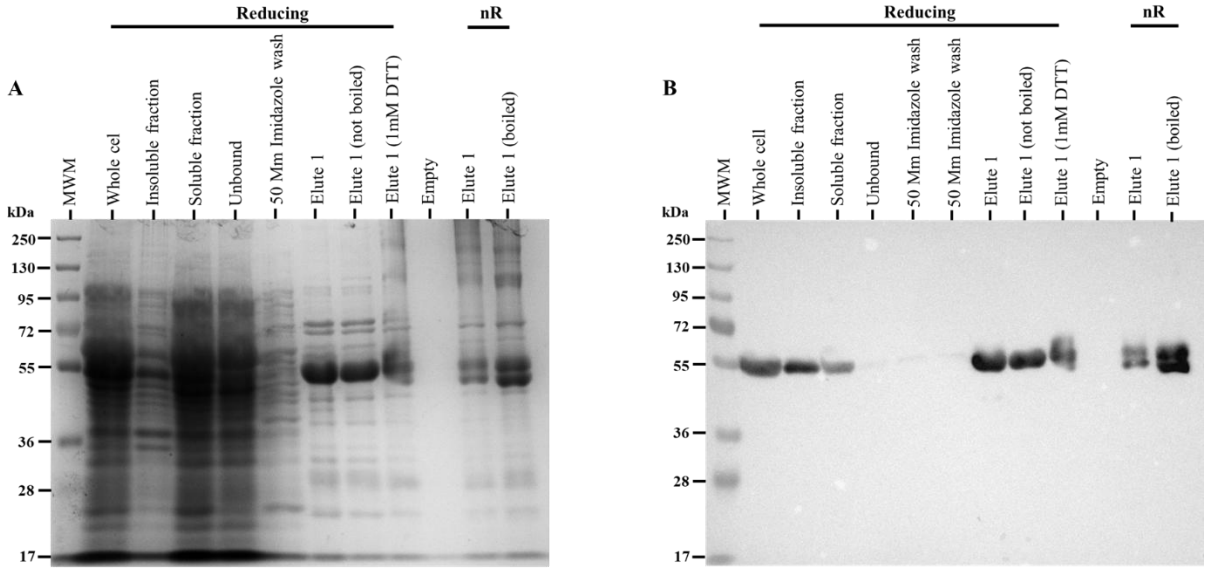

Figure S1 Analysis of recombinant PafA purification by IMAC Samples from the Ni^{2+}-NTA purification were resolved on 10% SDS-PAGE gels under both reducing and non-reducing (nR) conditions. (A) Coomassie Brilliant Blue staining reveals the total protein profile. (B) Immunoblot analysis using a mouse monoclonal anti-6xHis primary antibody and an HRP-conjugated goat anti-mouse secondary antibody, with chemiluminescent (ECL) detection.

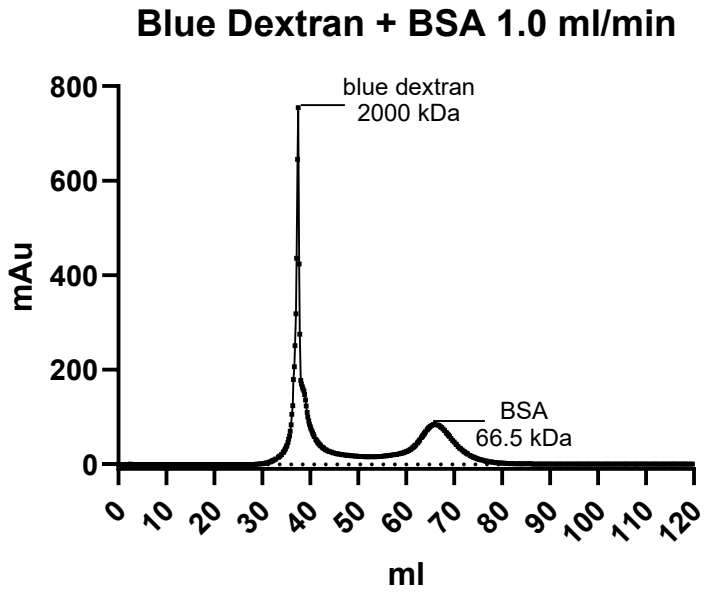

Figure S2 Size-exclusion chromatography calibration with protein standards. The elution profile from a Sephacryl S200 column for the molecular weight standards Blue Dextran 2000 (~2,000 kDa; void volume, 37.5 ± 0.03 mL, n=4) and bovine serum albumin (BSA, 66.5 kDa; 65.0 ± 0.28 mL, n=4). Standards were prepared at 2.5 mg/mL and 5 mg/mL, respectively, and chromatographed at a flow rate of 1 mL/min.

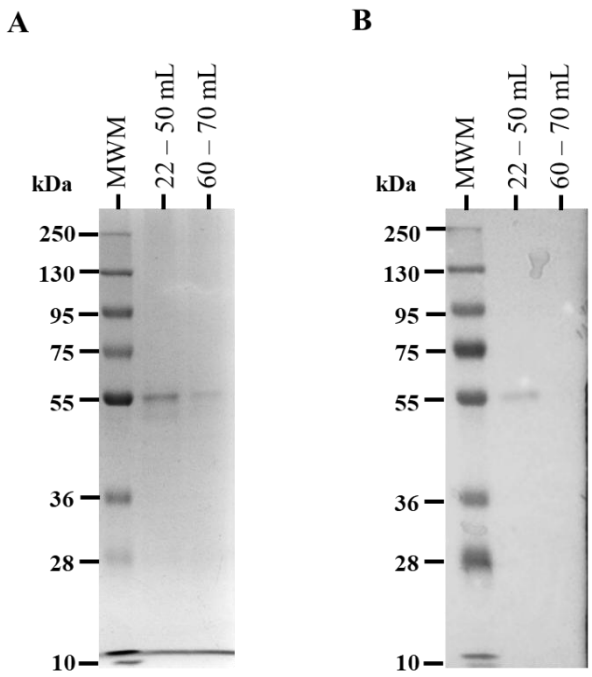

Figure S3 Recombinant PafA SEC purification. Sephacryl S-200 SEC purification samples were pooled and concentrated using 10 kDa centrifugal filters. The concentrates were then resolved on 10% reducing SDS-PAGE gels. (**A**) Coomassie Brilliant Blue staining reveals the total protein profile. (**B**) Immunoblot analysis using a mouse monoclonal anti-6xHis primary antibody and an HRP-conjugated goat anti-mouse secondary antibody, with ECL detection.

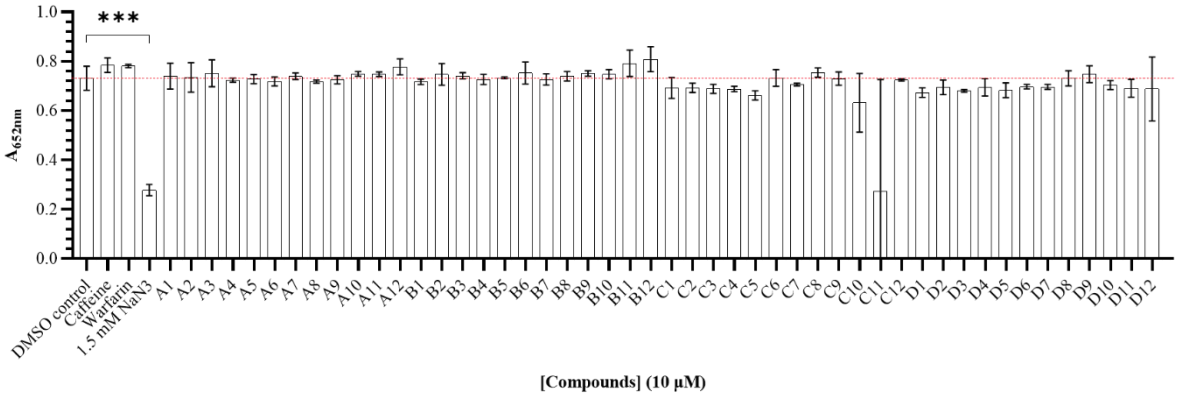

Figure S4 Screening for non-specific inhibitors via an optimised horseradish peroxidase non-specific inhibitor assay. A subset of compounds from the DiverSET CHEMBRIDGE library were tested at 10 µM. Compounds were pre-incubated with 100 ng/mL HRP in PBS (pH 7.4) for 5 minutes, and the reaction was initiated by adding 50 ng/mL TMB substrate containing 0.02% H_2O_2. NaN_3 served as the positive control, including the negative controls caffeine and warfarin. Absorbance at 652 nm was measured after 10 minutes reaction initiation at room temperature (n=3). Data are presented as mean ± standard deviation. Statistical significance was determined via an unpaired Welch's t-test (***$p < 0.001$)

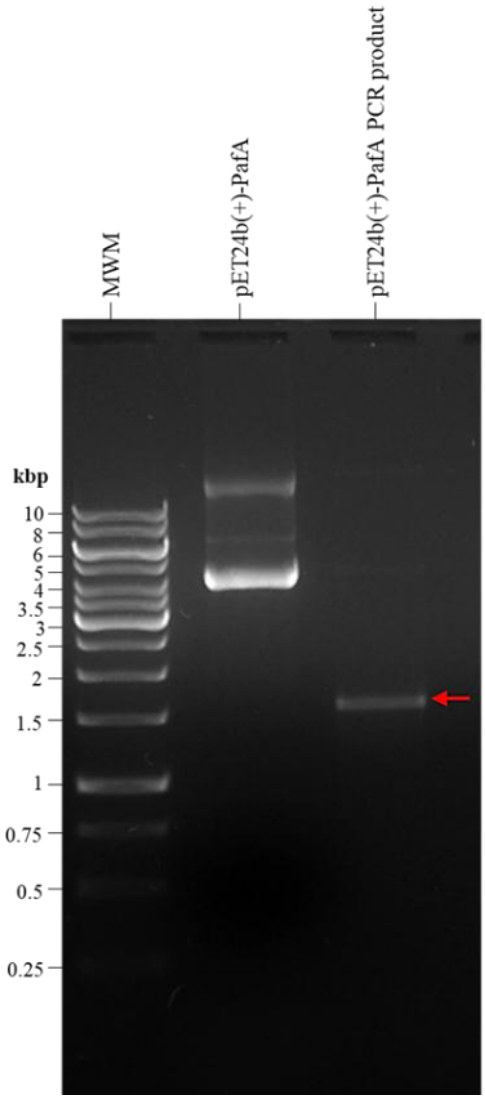

Figure S5. PCR amplification of the *M. tuberculosis* (H37Rv) *pafA* gene from the pET24b(+)-PafA vector construct. Vector DNA (10 ng) was aseptically added to a PCR mixture containing gene-specific primers and 1× *Taq* polymerase master mix. Following amplification, reactions were resolved on a 1% (w/v) agarose gel stained with 0.005% (v/v) ethidium bromide and visualised under ultraviolet light. **Lane MWM:** O'GeneRuler 1 kbp DNA ladder (Thermo Fisher Scientific, Massachusetts, USA). **Lanes 2 and 3:** recombinant vector DNA and PCR product migration, respectively. The red arrow indicates the migrated PCR product, which was confirmed to have 100% sequence identity by Sanger sequencing. The amplified product migrates at approximately 1.7 kbp, which is larger than the expected 1.4 kbp due to the inclusion of vector-derived sequences flanking the *pafA* gene (Rf).

Tool used in Schrödinger drug discovery suite: Protein secondary alignment

‘SSA’ denotes secondary structure alignment with each letter representing a secondary structure: C – chain (like an area that aligns but is not an alpha-helix or a beta-sheet), E – beta-sheet and H – alpha-helix

Score at the bottom

Structural alignment:

```
                .....+.....+.....+.....
cgPAfA   SSA 32    3 CCCCCCCCCEEEEEEEEEEEEECCEEECCHHHHHHHHHH
mtbPafA  SSA 3     1 -----CCCCEEEEEEEEEEEECCECCCCHHHHHHHHHC
cgPAfA          32    3 TVESALTRRIMGIETEYGLTFVDGDSKKLRPDEIARRMFR
mtbPafA          3     1 -----MQRRIMGIETEFGVTCTFHGHRRLSPDEVARYLFR

                .....+.....+.....+.....
cgPAfA   SSA 32   43 HCCCCCCCEEEEECCCEEEEEECCCEEEEEECCECCHHHH
mtbPafA  SSA 3    36 HHHHHHCCEEEEECCCEEEEEECCCEEEEEECCECCHHHH
cgPAfA          32   43 PIVEKYSSSNIFIPNGSRLYLNVGSHPEYATAECDNLTQL
mtbPafA          3    36 RVVSWGRSSNVFLRNGARLYLDVGSHPEYATAECDSLVQL

                .....+.....+.....+.....
cgPAfA   SSA 32   83 HHHHHHHHHHHHHHHHHHHHHHHHHHCCCCEEEEECCCCCC
mtbPafA  SSA 3    76 HHHHHHHHHHHHHHHHHHHHHHHHHHCCCCEEEEECCCCCC
cgPAfA          32   83 INFEKAGDVIADRMAVDAEESLAKEDIAGQVYLFKNNVDS
mtbPafA          3    76 VTHDRAGEWVLEDLLVDAEQRLADEGIGGDIYLFKNNTDS

                .....+.....+.....+.....
cgPAfA   SSA 32  123 CCECCEEEEEEECCCCCCHHHH-HHHHHHHHHHHCCCCCC
mtbPafA  SSA 3   116 CCECCEEEEEEEECCCCCHHHHHH-HHHHHHHHHCCCCCC
cgPAfA          32  123 VGNSYGCHENYLVGRSMPLKAL-GKRLMPFLITRQLICGA
mtbPafA          3   116 AGNSYGCHENYLIVRAGEFSRISD-VLLPFLVTRQLICGA

                .....+.....+.....+.....
cgPAfA   SSA 32  162 CEEECCCCCCCCEEECCCHHHHHHHHHHH------CCCCE
```

```
mtbPafA  SSA 3   155 CEEEEEC--CEEEEECCHHHHHH-CCE--EECCCCCCCCE
cgPAfA          32  162 GRIHHPNPSFPLGYCISQRSDHVWEGVSS------ASRPI
mtbPafA         3   155 GKVLQTP--KAATYCLSQRAEHI-WEG--VSSATTRSRPI

                    .........+.........+.........+.........

cgPAfA   SSA 32  204 EECCC-CCCCCCCCCCCEEECCCCCCCCHHHHHHHHHHHHH
mtbPafA  SSA 3   190 EE-CCCCCCCCCCCCEEEEECCCCCCCCHHHHHHHHHHHHH
cgPAfA          32  204 INTRD-EPHADSHSYRRLHVIVGDANMAEPSIALKVGSTL
mtbPafA         3   190 IN-TRDEPHADAEKYRRLHVIVGDSNMSETTTMLKVGTAA

                    .........+.........+.........+.........

cgPAfA   SSA 32  243 HHHHHHHHCCCCCCCCCCCHHHHHCHHHHHCCCCEEEEEE
mtbPafA  SSA 3   229 HHHHHHHHCCCCCCCCCCCCHHHHHHHHHHHCCCCCCEEEE
cgPAfA          32  243 LVLEMIEADFGLPSLELANDIASIREISRDATGSTLLSLK
mtbPafA         3   229 LVLEMIESGVAFRDFSLDNPIRAIREVSHDVTGRRPVRLA

                    .........+.........+.........+.........

cgPAfA   SSA 32  283 CCEEEEHHHHHHHHHHHHHHHHHHHHCCCCCCCCHHHHHH
mtbPafA  SSA 3   269 CCEEECHHHHHHHHHHHHHHHHHHCC-----C--HHHHH
cgPAfA       32  283 DGTTMTALQIQQVVFEHASKWLEQRPEPEFSGTSNTEMAR
mtbPafA      3   269 GGRQASALDIQREYYTRAVEHLQTREP-----N--AQIEQ

                    .........+.........+.........+.........

cgPAfA   SSA 32  323 HHHHHHHHHHHHHHHCCCCCCCCCCHHHHHHHHHHHHHHHH
mtbPafA  SSA 3   302 HHHHHHHHHHHHHHHCCCCCCCCCCHHHHHHHHHHHHHHHH
cgPAfA          32  323 VLDLWGRMLKAIESGDFSEVDTEIDWVIKKKLIDRFIQRG
mtbPafA         3   302 VVDLWGRQLDAVESQDFAKVDTEIDWVIKRKLFQRYQDRY

                    .........+.........+.........+.........

cgPAfA   SSA 32  363 CCCCCCHHHHHHHHHHHHCCCCCCHHHHHHHHCCCCCCCC
mtbPafA  SSA 3   342 CCCCCCHHHHHHHHHHHHCCCCCCHHHHHHHHCCCCCCCC
```

```
cgPAfA          32  363 NLGLDDPKLAQVDLTYHDIRPGRGLFSVLQSRGMIKRWTT
mtbPafA          3   342 DMELSHPKIAQLDLAYHDIKRGRGIFDLLQRKGLAARVTT


                  .........+.........+.........+.........
cgPAfA   SSA 32  403 HHHHHHHHHHHCCCCHHHHHHHHHHHHHHHHHHCCEEEEECEE
mtbPafA  SSA 3   382 HHHHHHHHHHHHCCCHHHHHHHHHHHHHHHHHHCCEEEEECEE
cgPAfA          32  403 DEAILAAVDTAPDTTRAHLRGRILKAADTLGVPVTVDWMR
mtbPafA          3   382 DEEIAEAVDQPPQTTRARLRGEFISAAQEAGRDFTVDWVH


                  .........+.........+.........+.........
cgPAfA   SSA 32  443 EEEECCCC-EEECECCCCCCCCHHHHHHHHHHHHHHH
mtbPafA  SSA 3   422 EEEEC-CCCEEECECCCCCCCCHHHHHHHHHC----
cgPAfA          32  443 HKVNRPEP-QSVELGDPFSAVNSEVDQLIEYMTVHA
mtbPafA          3   422 LKLND-QAQRTVLCKDPFRAVDERVKRLIASM----

Alignment Score:  0.066 (smaller is better)
        RMSD:  1.263 Angstrom
```