Understanding Enzymatic Mechanism and Allostery Using Macromolecular Rate Theory

A thesis

submitted in fulfilment

of the requirements for the degree

of

Doctor of Philosophy in Biology

at

The University of Waikato

by

Emily Katherine Grout

THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

2022

Abstract

Intricate systems of regulation at many levels act to control enzyme rates to tune metabolism across biosynthetic pathways. Allostery is a ubiquitous regulatory mechanism for feedback and regulation of enzymes in biosynthetic pathways. ATP-phosphoribosyltransferase from *Mycobacterium tuberculosis* (*mtu*ATP-PRT) catalyses the first committed step in *de novo* histidine biosynthesis, and allosteric product inhibition of ATP-PRT by L-His is key to regulating metabolic flux in this pathway. *mtu*ATP-PRT also exhibits allosteric activation by a histidine analogue 3 (2-Thienyl)-L-alanine (TIH). This compound binds in the same allosteric site as L-His and appears to elicit a structurally identical "tensed" conformation, apparently contradicting previous suggestions that the tensed conformation drives inhibitory regulation in *mtu*ATP-PRT.

Through mutations adjacent to the allosteric binding site, we have interrupted allosteric inhibition by L-His while retaining allosteric activation by TIH. Kinetic assays of wild type ATP-PRT and mutants across a large temperature range with various allosteric effects, has identified dynamic changes between allosterically modulated states of the enzyme, and suggests an enthalpy-entropy ($\Delta H^{\ddagger}$-$\Delta S^{\ddagger}$) trade off occurring in allosteric states.

Complementing our kinetic data, crystallographic analysis and molecular dynamics (MD) simulations of wild type and mutant ATP-PRT have identified two residues that differ upon the binding of TIH (in comparison to L-His binding). Glu-18 and Arg-27 rotate between ligand bound states and form interacting bonds in L-His bound structures that may serve to stabilise and reduce the overall flexibility of the hexameric conformation. This Glu-18 and Arg-27 interaction varies between L-His-bound, APO-*mtu*ATP-PRT and TIH-bound structures, suggesting this interaction may demarcate allosteric activation and inhibition.

An evolutionary analysis of ATP-PRT was undertaken and has identified the long form of ATP-PRT as the ancestral form, and the short form ATP-PRT has subsequently co-opted a histidyl-tRNA synthetase (HisRS) gene product (HisZ) as its regulatory domain. Aminoacyl-tRNA synthetases are known to be prone to

horizontal gene transfer, therefore the co-option of HisRS and the loss of the C-terminal domain in ATP-PRT$_S$ provides an evolutionary trajectory for diverse ATP-PRT isoforms.

Acknowledgements

Firstly, I would like to thank my supervisors Professor Vic Arcus and Professor Emily Parker for this opportunity. Thank you both for your incredible *incredible* expertise, guidance, and enthusiasm about MMRT and ATP-PRT.

Thank you to Dr Erica Prentice for providing me with the skills, lab support, and teaching me all the tips and tricks around the protein lab. Thank you to Dr Gerd Mittelstädt for helping me from afar! Thank you for sharing your expertise and magic fixes for ATP-PRT, and for generously collecting ITC data. To both Erica and Gerd I don't know how far I would have gotten without your training and ever helpful suggestions.

Dr Wanting Jiao, I am so appreciative of the simulations you were able to run from a distance. While I am immensely disappointed we were unable to meet up for collaboration as proposed I'm grateful for the Zoom meetings and emails shared to discuss those fantastic results.

Thank you to Erica, Elizabeth Rzoska-Smith, Keely Oldham, Carlin Hamill, Jack McGarvie, and Andrew Howard for taking your time to read chapters in this thesis. To everybody else in the C.2 lab (past and present), you have made the lab such an incredible and fun place to work. Your support and banter have made this PhD journey all the better, even when nothing was working, or we were stuck at home during the pandemic. Keely and Carlin I will always be grateful for the crystallography crash courses, your companionship during late night data collection, and your willingness to join me in fixing and fiddling with equipment to get it just right.

I would like to thank the University of Waikato for providing me with a Doctoral Scholarship that allowed me to undertake my research over the last three years. Also, a big thank you to the Waikato Graduate Women's Education Trust for the Doctoral Study Award.

Finally, thank you to my family, and especially Andrew for your incredible support during my PhD. Thank you all for taking the time to listen, and politely nod along during periods of excitement and frustration. I appreciate all of the support and words of wisdom I received – but sorry Dad, I can't run an electrical current through the spec while running assays.

Table of Contents

List of Figures

List of Tables

List of Equations

List of Abbreviations

aaRSs	Aminoacyl-tRNA synthetases
Abs	Absorbance
ACT	Aspartate kinase, chorismate mutase and TyrA
AICAR	5-Aminoimidazole-4-carboxamide ribonucleotide
AICc	Akaike information criterion
APS	Ammonium persulfate
ATP	Adenosine-5' triphosphate
ATP-PRT	Adenosine triphosphate phosphoribosyltransferase
*cgl*ATP-PRT	*Corynebacterium glutamicum* ATP-PRT
*cje*ATP-PRT	*Campylobacter jejuni* ATP-PRT
*C*p	Heat capacity at constant pressure
C-terminal	Carboxyl-terminus
Da	Daltons
E_A	Activation energy
EC	Enzyme commission number
*Eco*PPase	*Escherichia coli* pyrophosphatase
g	Times the force of gravity
H	Enthalpy
HisG_C	C-terminal regulatory domain of ATP-PRT_L
$hisG_L$	ATP-PRT_L gene
$HisG_L$	ATP-PRT_L protein
$hisG_S$	ATP-PRT_S gene
$HisG_S$	ATP-PRT_S protein
His-tag	Poly-histidine tag
IMAC	Immobilised metal affinity chromatography
IPA	Isopropyl alcohol
IPTG	Isopropylthio-β-D-galactosidase
ITC	Isothermal titration calorimetry
k	Rate constant
Kb	Kilobase
k_B	Botzmann constant
k_{cat}	Enzyme rate constant
K_D	Dissociation constant
kDa	Kilo Dalton
K_i	Inhibition constant
K_M	substrate binding constant
LB	Luria Bertani media
L-His	L-Histidine
MD	Molecular dynamics
MM	Michaelis-Menten
MMRT	Macromolecular rate theory
*mtu*ATP-PRT	*Mycobacterium tuberculosis* ATP-PRT
MW	Molecular weight
OD_{600}	Optical density at 600 nm
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction

PDB	Protein Data Bank
PP_i	Di-phosphate
PRPP	Phosphoribosyl pyrophosphate
RMSD	Root mean square deviation
s	Second
S	Entropy
SD	Standard deviation
SEC	Size exclusion chromatography
T	Temperature
T_0	Reference temperature
TEMED	N, N, N, N,-tetramethylethylenediamine
TEV	Tobacco Etch Virus Protease
TIH	3-(2-Thienyl)-L-alanine
T_{inf}	Temperature inflection point
T_m	Temperature midpoint for unfolding
T_{opt}	Temperature optimum for activity
TS	Transition state
v/v	Volume per volume
w/v	Weight per volume
WT	Wildtype

Chapter 1
Introduction

1.1 Introduction

Enzymes catalyse nearly all chemical reactions within cells and can dramatically increase chemical reaction rates (Fersht, 1999; Wolfenden & Snider, 2001; Daniel & Danson, 2010). Their catalytic effects impact a diverse number of reactions, with each enzyme primarily catalysing a single reaction due to their binding specificity. Enzymes work in concert, placed in order to form biosynthetic pathways that play essential roles in central metabolism, energy production, and *de novo* synthesis of essential metabolites. Intricate systems of regulation act at many levels to control enzyme rates, to tune metabolism across and between these biosynthetic pathways.

ATP-Phosphoribosyltransferase (ATP-PRT), the first committed enzyme step of histidine synthesis, initiates a metabolically expensive biochemical pathway, and is subject to sophisticated regulatory mechanisms including allostery. Regulation of ATP-PRT initiates histidine biosynthesis only when the amino acid is needed, and there is sufficient energy in the cell for histidine biosynthesis. The mechanism in which ATP-PRT functions, and the mechanism behind allosteric regulation in general is key to understanding the strategies of biological regulation in histidine biosynthesis and indeed other enzymes. It is fundamental to understand the normal function of enzymes, gaining a full understanding of what is happening in normal circumstances that allows us to exploit functions of enzymes through biochemical engineering for numerous applications, including drug design.

Macromolecular rate theory (MMRT) has been used to describe the curvature of enzyme rates with temperature (i.e. deviations from Arrhenius behaviour) and was first presented by Hobbs *et al.* (2013). Since the establishment of MMRT it has been used in the elucidation of biological mechanisms including temperature adaptation, growth rates, and ecosystem level respiration rates (Hobbs *et al.*, 2013; Schipper *et al.*, 2014; Arcus *et al.*, 2016; Prentice *et al.*, 2020). Applying MMRT

in the context of allosteric regulation may help elucidate the thermodynamic changes and mechanisms involved in allosteric regulation of ATP-PRT.

1.2 Allostery

Allostery is a mechanism for feedback and enzyme regulation. Allostery was named in 1961 by Jacques Monod and Francois Jacob as a term to describe inhibition of enzymes where the inhibitory molecule (bearing no similarity to substrate) bound to a distinct site, distal to the active site of the enzyme and exerts an alteration to enzyme activity (Motlagh *et al.*, 2014; Liu & Nussinov, 2016). Although not defined until 1961, the phenomenon of allostery was first observed by Christian Bohr in 1904 for haemoglobin. Carbon dioxide was found to affect the binding of oxygen to haemoglobin proteins, in that when carbon dioxide levels increased, the binding affinity of oxygen decreased (Liu & Nussinov, 2016).

Two models have been proposed to describe allosteric mechanism; The Monod-Wyman-Changeux (MWC) concerted mechanism (Monod *et al.*, 1965) and The Koshland-Némenthy-Filmer (KNF) sequential model (Koshland *et al.*, 1966). Both MWC and KNF attribute allosteric modulation to the conformational change of an enzyme, however, more recently X-ray crystallography and computational simulations have expanded these binary mechanisms allowing for a more dynamic view of allostery (Liu & Nussinov, 2016).

Allostery is a wide spread phenomenon and a critically important step in cell signalling (Nussinov *et al.*, 2013). Allosteric feedback is a highly sophisticated regulatory system, commonly seen in oligomers involved in cell signalling and the regulation of metabolism, and is important in regulating energy flux within the cell. These oligomeric proteins often exist as even numbered multimeric proteins (i.e: dimers, tetramers, hexamers, *etc.*) offering even symmetry to the enzyme (Li *et al.*, 2013).

1.2.1 Models of Allosteric regulation

The two primary models of allosteric regulation are binary conformational models (MWC and KNF). However, more recently the morpheein model, ensemble

mechanisms, and allosteric models related to protein dynamics have been discussed to further explain allostery and its effect on the conformational landscape (Cooper & Dryden, 1984; Jaffe, 2005).

1.2.1.1 Concerted (Monod-Wyman-Changeux)

MWC was proposed based on the structural details of haemoglobin and the model contains many generalised assumptions about protein structure (Changeux, 2012). MWC (**Figure 1.1**) relies on two principles; that allosteric regulation exists in oligomers, and that there is an equilibrium between two differently stabilised structural confirmations – tensed (**T**) and relaxed (**R**) states (Fersht, 1999; Changeux, 2012). Also dubbed the symmetry model (as each subunit adopts the same **T** or **R** state), MWC assumes each oligomer has identical subunits. Each of these subunits can either exist in **T** or **R** states independently of allosteric modulation (Fersht, 1999). In the MWC model, the binding of an allosteric ligand triggers a shift of the equilibrium, in the instance of allosteric inhibition this would shift the equilibrium to favour the **T** state, and *vice versa* for an agonist or activator. Following MWC, allosteric regulators function by shifting the equilibrium of **T** and **R** states, thus altering the distribution of **T** and **R** states. The MWC model makes the assumption that allosteric enzymes contain internal symmetry within the oligomer structure, and that **T** and **R** states are rotations of the monomer unit (Cui & Karplus, 2008).

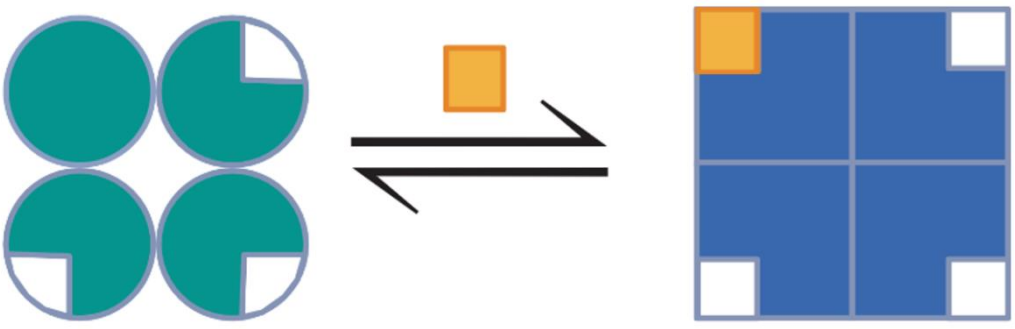

Figure 1.1. Schematic diagram of Monod-Wyman-Changeux model of allosteric regulation (Monod *et al.*, 1965). Symmetrical oligomers exist in equilibrium between Relaxed (green) and Tensed (blue) states. Effector molecules (orange) shift the equilibrium between states. Adapted from (Fersht, 1999).

Allosteric enzymes are noted to produce sigmoidal plots of substrate concentration versus rates (**Figure 1.2**, B), and they do not follow typical Michaelis-Menten kinetics due to multiple active sites with cooperativity between sites. The sigmoidal rate vs substrate concentration curve could be attributed to partial substrate binding in the **R** state as the equilibria shifts allowing for increased substrate binding at higher concentrations. (Fersht, 1999).

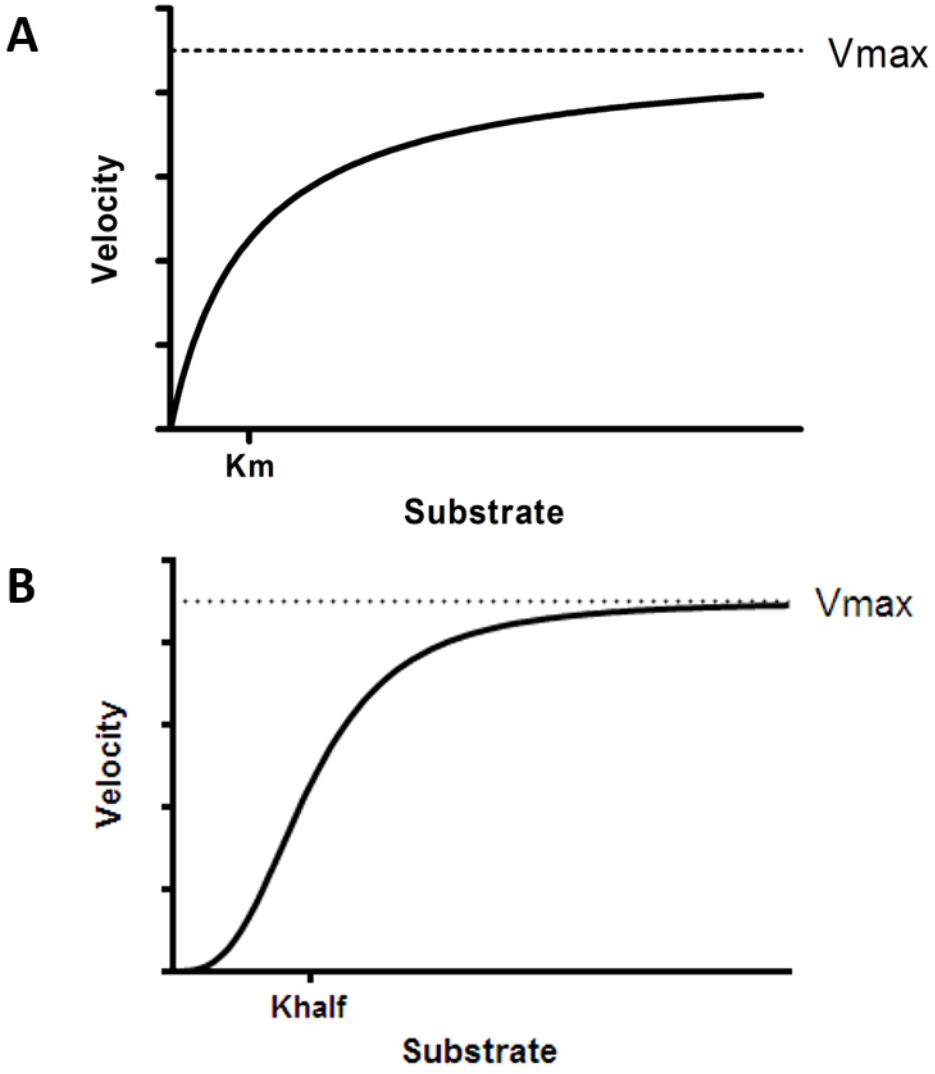

Figure 1.2. Enzyme kinetic schematic. Allosteric enzymes following MWC do not exhibit traditional Michaelis-Menten kinetics of a hyperbolic increase in rate in response to substrate concentration (A) but instead show sigmoidal rate curves (B) when cooperativity exists between enzyme subunits.

1.2.1.2 Sequential (Koshland-Némenthy-Filmer)

Like MWC, the KNF model of allosteric control assumes enzymes exist in **T** or **R** states and are multimeric. Also called the induced fit theory, KNF describes a level of cooperativity between protein subunits of the oligomer. Unlike MWC, KNF proposed the structural conformation of the protein would change upon the binding of an allosteric ligand, therefore altering the proteins affinity for a ligand at each other binding site (**Figure 1.3**). In comparison to MWC, the KNF model proposes that each subunit of an oligomer changes independently, however allosteric ligand binding in one subunit would have a universal effect thereby increasing the likelihood of other subunits to adopt the same conformational change sequentially. Like MWC, upon the binding of an allosteric ligand to the oligomer, the equilibrium would shift towards either the **T** or **R** state (Koshland *et al.*, 1966; Fersht, 1999).

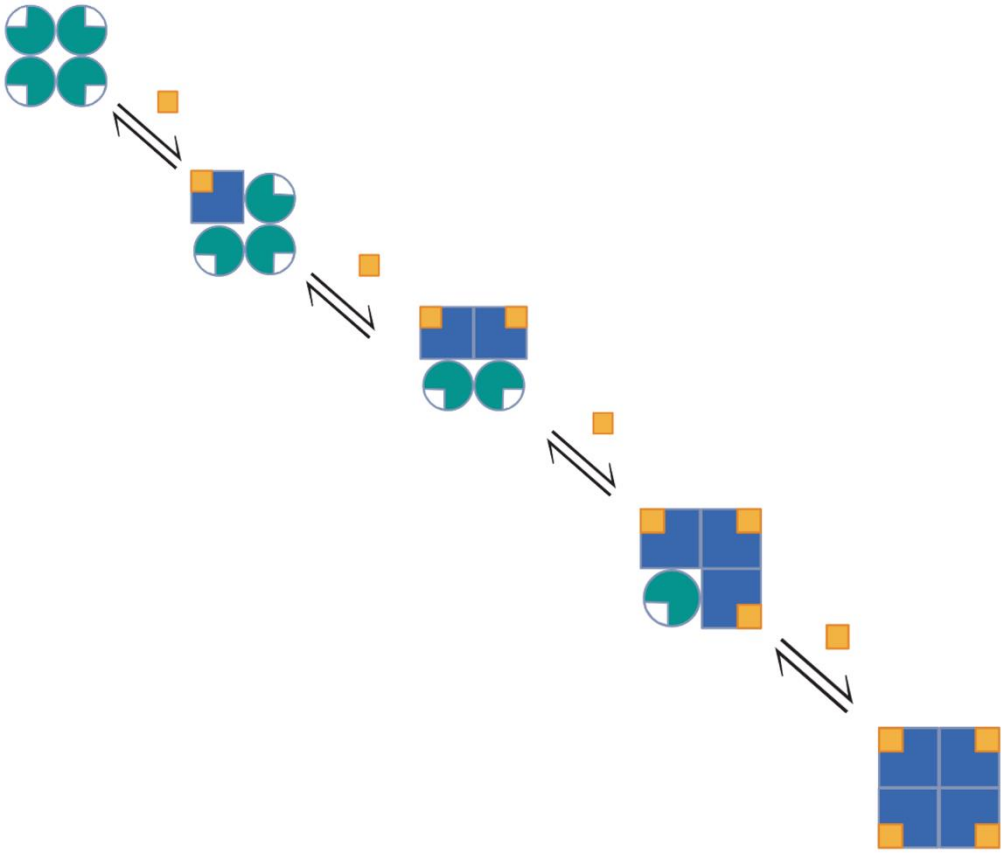

Figure 1.3. 2D schematic of Koshland-Némenthy-Filmer (KNF) model of allosteric regulation. Symmetrical oligomers exist in equilibrium between Relaxed (green) and Tensed (blue) states. Effector molecules (orange) shift the equilibrium between states. In the KNF model each protein subunit can alter conformation independently, thus creating the 'induced fit' model. Figure adapted from (Fersht, 1999).

KNF also differs from MWC in that KNF model can induce negative cooperativity, structural changes that decrease ligand binding compared to a population shift towards **R** states (Fersht, 1999).

1.2.1.3 Stereochemical Model

In 1970 Max Perutz incorporated X-ray crystallography and structural biology into the stereochemical allosteric model using the first crystal structure of haemoglobin (Perutz *et al.*, 1960; Perutz, 1970). Structural analysis of haemoglobin identified salt bridges near the allosteric sites that could mediate conformational changes upon oxygen binding. MWC and KNF omit structural details that contribute to allosteric signal transduction *i.e.*, how allosteric communication is achieved between the allosteric binding site and active site. Residue-residue interactions have been described by Motlagh et al. (2014) and del Sol, Tsai, Ma, and Nussinov (2009) that serve to transmit signals from allosteric sites to the active site of an enzyme (**Figure 1.4**), defining the details of the tensed and relaxed states.

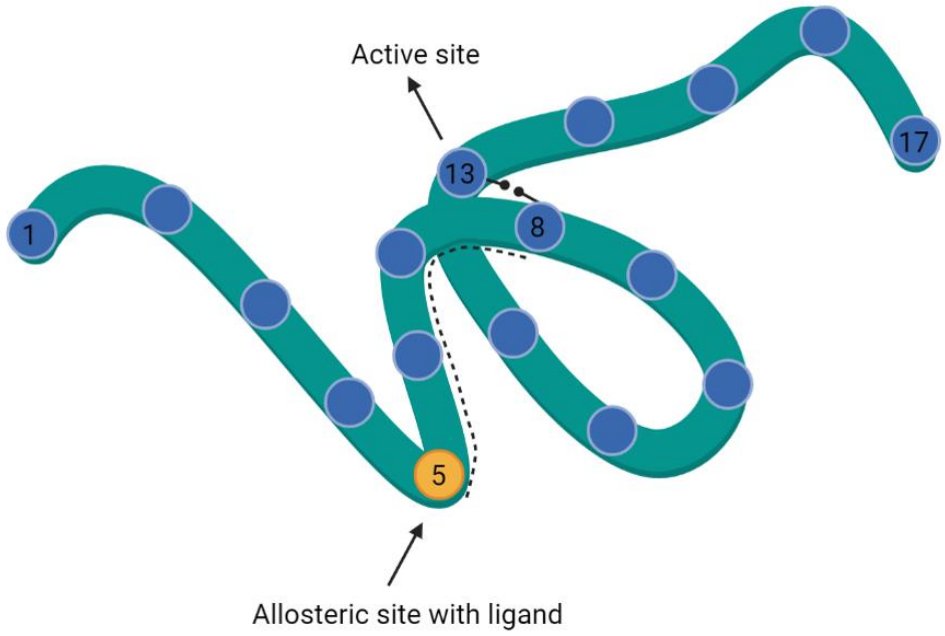

Figure 1.4. 2D diagram of the stereochemical model of allosteric regulation. A simplified 17 residue long protein chain is shown in green. A pathway of allosteric signal transmission is drawn from residue five (orange; allosteric site) to residue eight. Dotted arrows indicate side chain interactions between residue eight and residue 13 of the active site. This figure illustrates the roles of residues five to eight and 13 in allosteric signalling. Mutations to these residues would alter allosteric function, however mutations to the loop (residues 9-

12) would have no effect on allosteric response as they do not form allosteric interactions. Figure adapted from (del Sol *et al.*, 2009).

1.2.1.4 Morpheein

The morpheein model, named after morpheein proteins, postulates that enzymes exist in a highly dynamic equilibrium of multiple oligomer states. First proposed by Jaffe (2005) the morpheein model has similarities to MWC and KNF, and is described as a concerted dissociative model. However, the major difference in the morpheein model is the suggestion that enzymes are able to rearrange into different oligomeric assemblies (*i.e*. trimer to tetramer), and the equilibria of oligomeric assemblies is driven by allosteric ligand binding (**Figure 1.5**).

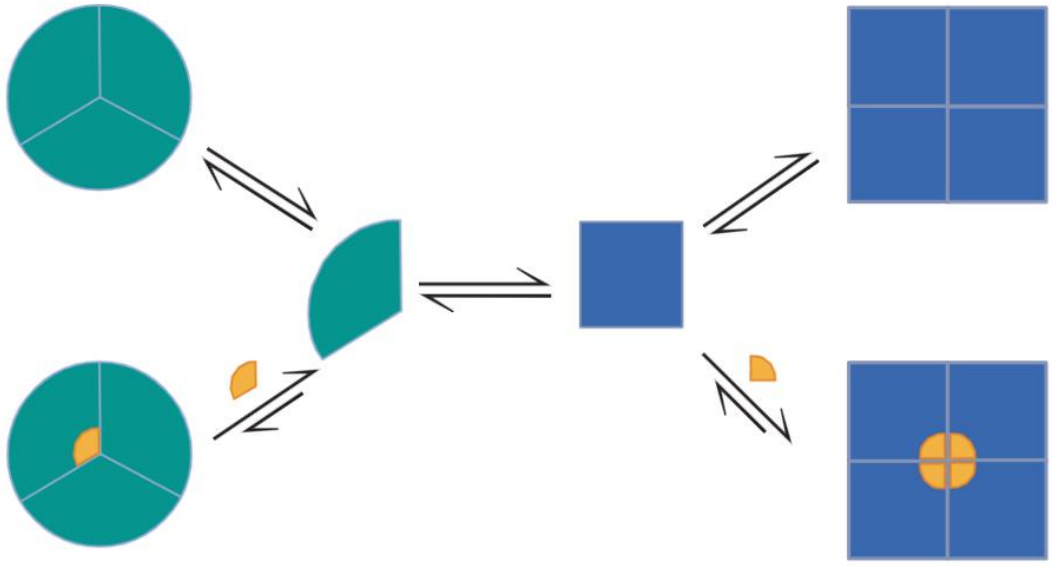

Figure 1.5. 2D schematic representation of the morpheein model of allosteric regulation. Homo-oligomeric allosteric proteins exist in equilibrium between multiple monomeric conformations with differing catalytic capacity (green circle fragment and blue square). These monomeric conformations can each rearrange into distinct oligomer assemblies (trimers or tetramers) and the binding of allosteric ligands drives the shift into one of these distinctive oligomeric assemblies. Figure adapted from (Jaffe, 2005).

Morpheein allosteric regulation was developed to describe the two distinct oligomer assemblies of porphobilinogen synthase (PBGS) (Jaffe, 2005). The dissociative allosteric mechanism in PBGS sees the enzyme dissociate from an octamer to a dimer, which then rearranges into a less active hexamer. When in the active octameric arrangement, Mg^{2+} stabilises a subunit interface, but this binding site is absent in the hexameric form (Jaffe, 2016). The morpheein model of allostery

entails dissociative movements of protein chains that are highly regulated to achieve reassembly of the quaternary structure. The morpheein model appears to be a special case of allosteric mechanism and is not often observed in other allosteric systems. However, the morpheein model in PBGS illustrates the highly dynamic and variable nature of allosteric mechanisms observed across enzyme systems (Jaffe, 2016).

1.2.1.5 Ensemble allosteric model

The ensemble model of allosteric regulation builds on MWC principals. In the ensemble model enzymes are treated as ensembles of **T** and **R** states that are able to interact with one another (Hilser *et al.*, 2012; Motlagh *et al.*, 2014). The ensemble model allows a great number of microstates of subunits to co-exist compared to classical MCW and KNF models.

Figure 1.6. 2D schematic model of the ensemble allosteric model, representing the range of states in equilibrium for a dimer. Protein subunits can exist harmoniously as tensed (square) or relaxed (circles). These subunits can independently exist as ligand bound (orange fill) or ligand unbound (no fill). Horizontal black lines between subunits indicates finite interaction energy. Figure adapted from (Hilser *et al.*, 2012).

1.2.1.6 Dynamic allostery

In 1984 Cooper and Dryden proposed that allosteric mechanisms have underlying dynamic origins instead of fixed conformational changes (Cooper & Dryden, 1984). Proteins could be viewed as dynamic structures that were capable of undergoing motion and change, and ligand binding in allosteric regulation could contribute to structural and entropic changes. Dynamic origins of allostery were further supported by Colombo *et al.* (1992) through investigation of haemoglobin allostery. An additional 60 bonds were shown to be present in haemoglobin in one allosteric state, binding an additional 60 water molecules into the structure in the transition from **T** to **R** sates.

Not all allosteric enzymes are oligomeric, and this discounts conformational based allosteric methods such as MWC, KNF, and morpheein (Li *et al.*, 2013). However, the dynamic allostery model offers a model for enzymes that do not undergo obvious changes to protein structure upon ligand binding. Instead, changes in catalysis due to allosteric regulation can be attributed to changes in the protein's dynamics, a factor that is not generally captured in crystallographic investigations (Tsai *et al.*, 2008; Tzeng & Kalodimos, 2009). In essence, dynamic allostery suggests ligand binding elicits changes in protein dynamics, and these changes may be driven by changes to the overall entropy of the protein, and not simply static conformational changes (Tsai *et al.*, 2008). Dynamic allostery can be viewed as an extension of the MWC and KNF models. The traditional MWC and KNF models postulate two conformational states for the enzyme. A natural extension may to be postulate 4, 8 or 10 different conformational states for the enzyme. The dynamics model simply postulates large numbers of conformational states for the enzyme and thus, these models sit on a continuum.

1.2.1.7 Population shift model

The population shift model is an extension of MWC model, but suggests both the **T** and **R** states are always present without allosteric ligands (Ravasio *et al.*, 2019). Weber (1972) identified that allosteric binding leads to a change in these states' stability, triggering a conformational change of the protein toward the favoured state. This population shift model has continued to grow as a model that combines

allosteric regulation and dynamic mechanisms that are increasingly associated with native enzyme activity (Lu *et al.*, 2014; Ravasio *et al.*, 2019).

Conformational models (MWC, KNF, and morpheein) are unable to model the high diversity of allosteric enzymes (Hilser *et al.*, 2012). Significant differences in allosteric enzyme oligomeric states, symmetry, activators and repressors cannot be explained using conformational models. MWC and KNF models also do not explain the underlying mechanism of allosteric communication and signal control like stereochemical, and dynamic models strive to achieve. Nonetheless, we can view all of these models as related and on a continuum from the symmetric binary MWC model, to the sequential binary KNF model, through to the recent fully dynamic models of allostery. The underlying fundamental question is the nature of the communication between the allosteric site and the active site that alters the conformations (in the case of MWC and KNF) or the dynamics of the system.

In the context of this fundamental question, ATP-PRT has been used as a model system to uncover the possible communication pathways between allosteric and active sites of an enzyme.

1.3 Adenosine triphosphate phosphoribosyl transferase in Histidine biosynthesis

Histidine biosynthesis is an ancient metabolic pathway that is also linked to a number of independent metabolic pathways in the cell. For these reasons, *de novo* histidine biosynthesis has been extensively studied in bacteria, archaea, lower eukaryotes, and plants (Fondi *et al.*, 2009; Kulis-Horn *et al.*, 2014).

Adenosine triphosphate phosphoribosyl transferase (ATP-PRT; EC. 2.4.2.17) catalyses the first committed step of histidine biosynthesis, converting adenosine triphosphate (ATP) and phosphoribosyl pyrophosphate (PRPP) into phosphoribosyl adenosine triphosphate (PR-ATP) (**Figure 1.7**). The conversion to PR-ATP is reversible, and is Mg^{2+} dependent (Ames *et al.*, 1961). *De novo* histidine synthesis is considered to require 41 ATP equivalents making it one of the most energetically demanding amino acids to synthesise (Brenner & Ames, 1971). Consequently,

ATP-PRT, as the entry point to this pathway, is subject to numerous levels of regulation. Competitive inhibition from adenosine monophosphate (AMP) and adenosine diphosphate (ADP), which signal low cellular energy, downregulate the enzyme (Morton & Parsons, 1977). Step five of histidine biosynthesis releases 5-Aminoimidazole-4-carboxamide ribonucleotide (AICAR), which then feeds into the purine biosynthesis pathway. AICAR has been identified as a competitive inhibitor to *Escherichia coli* ATP-PRT (*eco*ATP-PRT), mimicking ATP binding (Malykh *et al.*, 2018). Of key interest to this research, ATP-PRT is allosterically inhibited by the end product of the pathway L-His. Intriguingly, despite all the regulatory mechanisms of ATP-PRT that exist to regulate energy flux, *Mycobacterium tuberculosis* ATP-PRT (*mtu*ATP-PRT) has also been shown to be allosterically activated by the L-His analogue 3-(2-Thienyl)-L-alanine (TIH) (Pisco *et al.*, 2017). This makes this enzyme an ideal model system to study the thermodynamics of allosteric activation and inhibition.

ATP-PRT catalyses with a bi-bi kinetic mechanism (Kleeman & Parsons, 1976; Morton & Parsons, 1976) with ATP binding first, then PRPP, and next chemical catalysis, followed by the release of diphosphate, then the release of PR-ATP product.

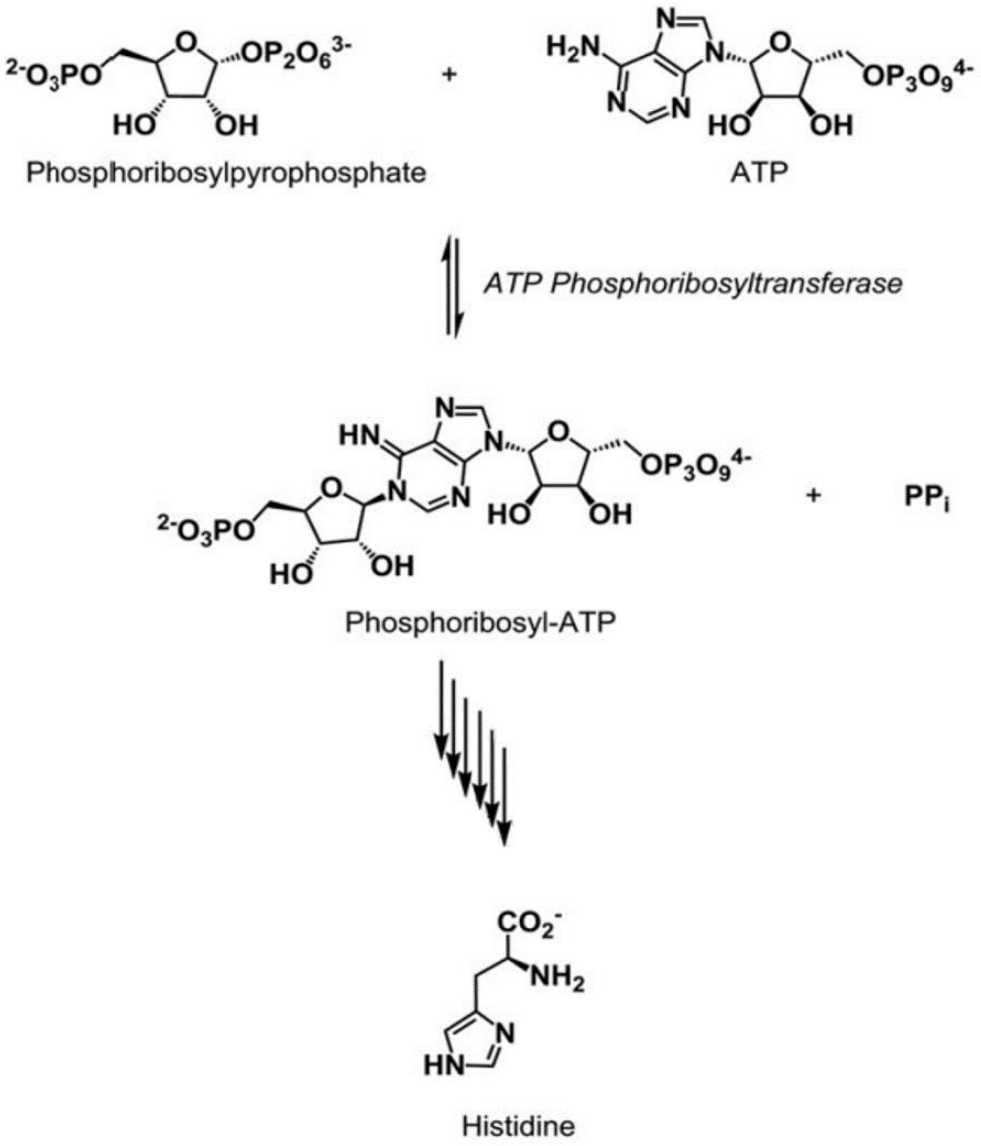

Figure 1.7. Abridged L-His biosynthesis pathway initiated by the reversible condensation of ATP and PRPP to PR-ATP by the enzyme ATP-PRT. Six additional enzymatic steps by four enzymes produce L-His. Figure adapted from (Mittelstädt *et al.*, 2016).

ATP-PRT is a rate limiting enzyme in histidine biosynthesis in all organisms. Due to its regulatory role, and absence in higher eukaryotes, it has been studied in a variety of bacterial species, both to understand histidine biosynthesis and assess its use as a drug target for human diseases (Martin, 1963; Vega *et al.*, 2005; Champagne *et al.*, 2006; Pedreño *et al.*, 2012; Zhang *et al.*, 2012; Mittelstädt *et al.*, 2016; Stroek *et al.*, 2017).

1.3.1 Diversity of ATP-PRT

ATP-PRT exists in two isoforms: long-form ATP-PRT (ATP-PRT$_L$) and short-form ATP-PRT (ATP-PRT$_S$). Both isoforms are active and catalyse the first committed step of histidine biosynthesis. Both isoforms are allosterically inhibited by the end product of the pathway they initiate (L-His). Although ATP-PRT$_L$ and ATP-PRT$_S$ share a near structurally-identical two domain catalytic core, they have disparate regulatory domains. Allosteric regulation is an efficient way of regulating an enzyme, and each proposed mechanism describes coordinated movements of domains that have arisen through evolution. To have two independent isoforms both exhibiting sophisticated allosteric regulation is impressive. Both isoforms are described below, and a detailed introduction and review of existing evolutionary models of ATP-PRT isoforms can be found in Chapter 2. A detailed structural review can be found in Chapter 5.

1.3.1.1 *hisG$_L$*

hisG$_L$ encodes the ~280-310 residues that form ATP-PRT$_L$ (**Figure 1.8)**. ATP-PRT$_L$ is generally a homo-hexamer consisting of three domains. Domains I and II form the catalytic core whilst domain III forms the regulatory subunit and allosteric binding site (Champagne *et al.*, 2006; Mittelstädt *et al.*, 2016).

1.3.1.2 *hisG$_S$*

hisG$_S$ encodes ~200 residues that form ATP-PRT$_S$ (**Figure 1.8**). These 200 residues form domains I and II, which function as the bi-lobal catalytic core, similar to that of ATP-PRT$_L$. Despite sharing the same structural catalytic core (domain I and II), the sequence identity between *hisG$_S$* and the h*isG$_L$* catalytic core is approximately 25 % (Champagne *et al.*, 2006). Unlike the long-form, ATP-PRT$_S$ generally exists as a hetero-octamer, forming a quaternary structure with *hisZ* (Livingstone *et al.*, 2016; Mittelstädt *et al.*, 2016). *hisZ* is derived from a group of histidyl-tRNA synthetase domains, and functions as the regulatory domain in ATP-PRT$_S$, offering enhanced catalysis and allosteric inhibition by L-His (Champagne *et al.*, 2006; Livingstone *et al.*, 2016).

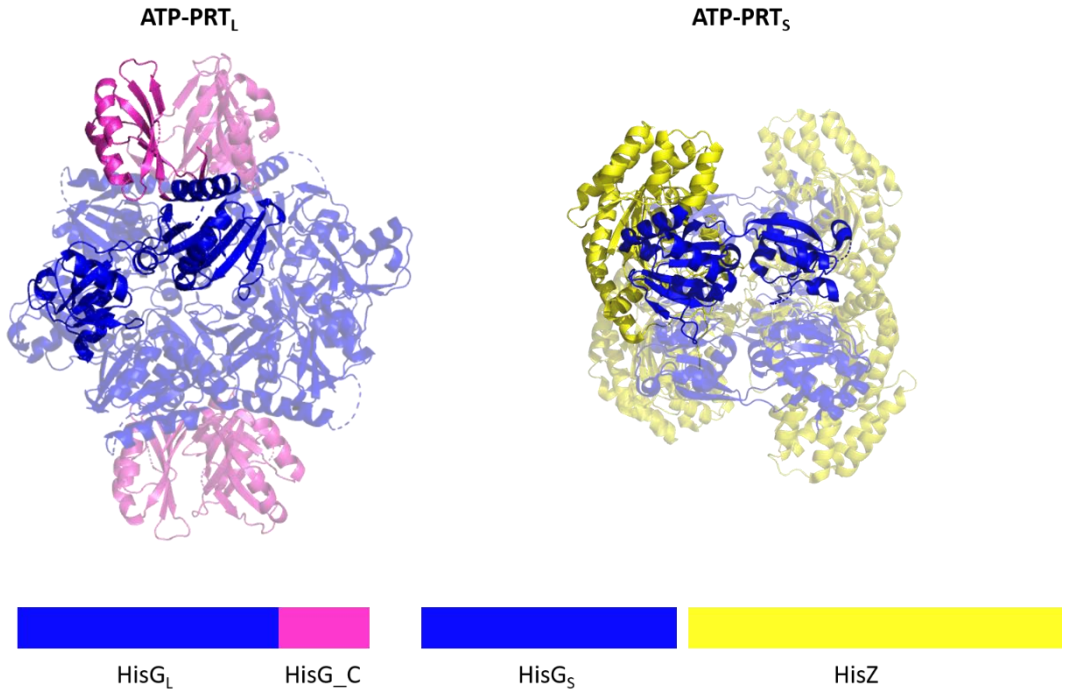

Figure 1.8. Gene and structure differences between ATP-PRT$_L$ and ATP-PRT$_S$. PDB 1NH7 (Cho *et al.*, 2003) illustrates homo-hexameric arrangement of ATP-PRT$_L$. This isoform contains a single gene product with a fused C-terminal regulatory domain encoded by *hisG_C*. ATP-PRT$_S$ arranges as a hetero-octamer formed from two gene products, HisG and HisZ. PDB 1Z7M (Champagne *et al.*, 2005) shows *hisG*$_S$ catalytic domains (blue) are arranged in pairs on either side of the tetrameric arrangement of HisZ regulatory domains (yellow) in the centre. The catalytic cores of both ATP-PRT isoforms organise as dimers in the quaternary structure. More details on gene productions and structural changes between ATP-PRT isoforms is described in Chapter 2.

1.3.2 Allosteric regulation of ATP-PRT

ATP-PRT from *Mycobacterium tuberculosis* (*mtu*ATP-PRT) is allosterically inhibited by L-His, the final product of the biochemical pathway ATP-PRT initiates. *mtu*ATP-PRT is a long-form ATP-PRT and thus contains a fused regulatory domain at the C-terminus. All ATP-PRT species with functional regulatory domains are allosterically inhibited by L-His. Additionally, *mtu*ATP-PRT has been reported by Pisco *et al.* (2017) to be allosterically activated by 3-(2-Theinyl-L-Alanine) (TIH) making *mtu*ATP-PRT an ideal enzyme of interest for investigations into allosteric mechanisms.

1.3.2.1 Allosteric inhibition by L-Histidine

Both ATP-PRT isoforms have been shown to be allosterically inhibited by histidine. (Martin, 1963; Vega *et al.*, 2005; Pisco *et al.*, 2017). This classic feedback inhibition sees ATP-PRT as a regulatory step in histidine biosynthesis. Allosteric inhibition by L-His has been deeply investigated in many ATP-PRT$_L$ species (Cho *et al.*, 2003; Pedreño *et al.*, 2012; Mittelstädt *et al.*, 2016; Mittelstädt *et al.*, 2018). In long-form ATP-PRT, allosteric inhibition by L-His elicits a conformational change to a closed form (**T** state) (**Figure 1.9**) and these have been observed in crystal structures from *Campylobacter jejuni* and *M. tuberculosis* (Cho *et al.*, 2003; Mittelstädt *et al.*, 2016; Pisco *et al.*, 2017). This **T** state is driven by a twist of the C-terminal regulatory domain likely due to the interactions extending from the histidine binding site to the domain interface (Cho *et al.*, 2003). A close view of L-His chemical structure, and ligand binding is shown in **Figure 1.10**.

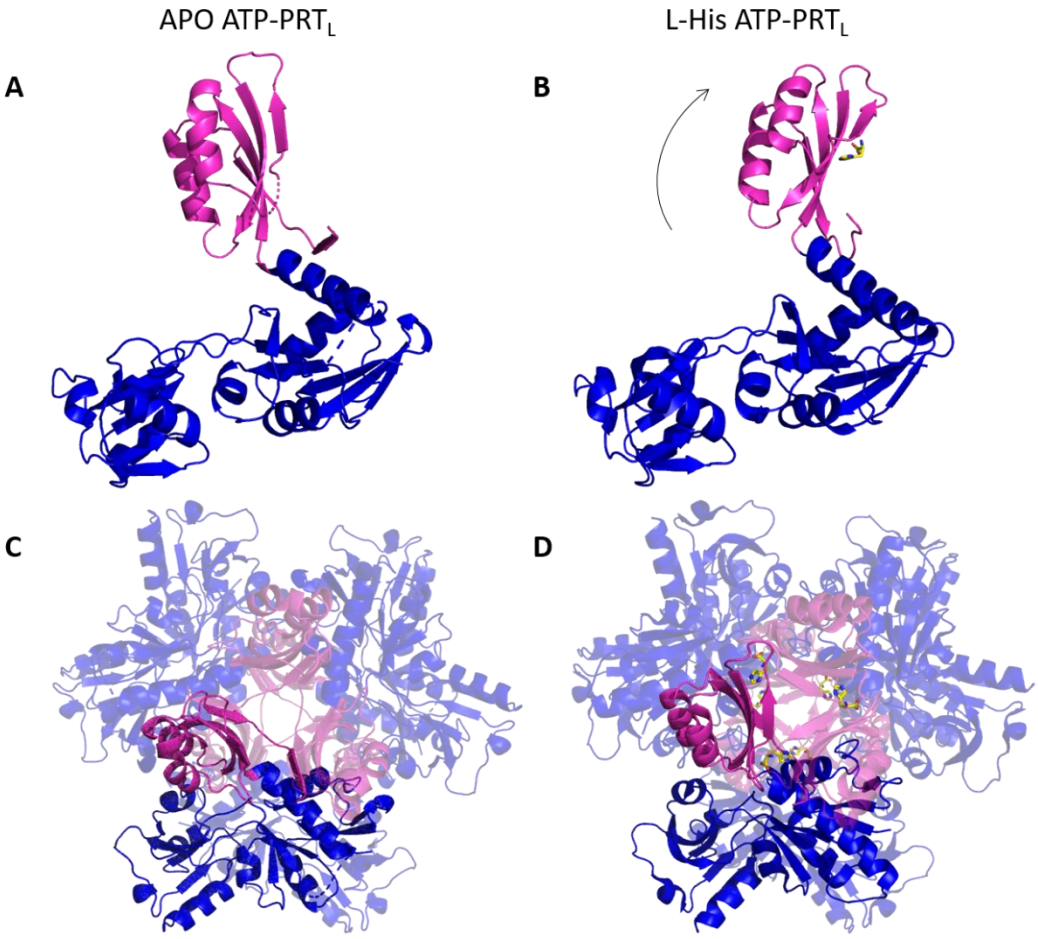

Figure 1.9. Tensed and relaxed states of *mtu*ATP-PRT between APO (A) and L-His bound (B) states. In the monomeric state L-His binding elicits the movement of the C-terminal regulatory domain (magenta) increasing the gap between domains I and III from 43 to 50

Å. This increased gap results in the tightening of the C-terminal trimeric interface (D), evident by the decreased space between β-sheets central to the hexamer, compared to APO hexamer state (C) which features a larger space at the trimeric interface.

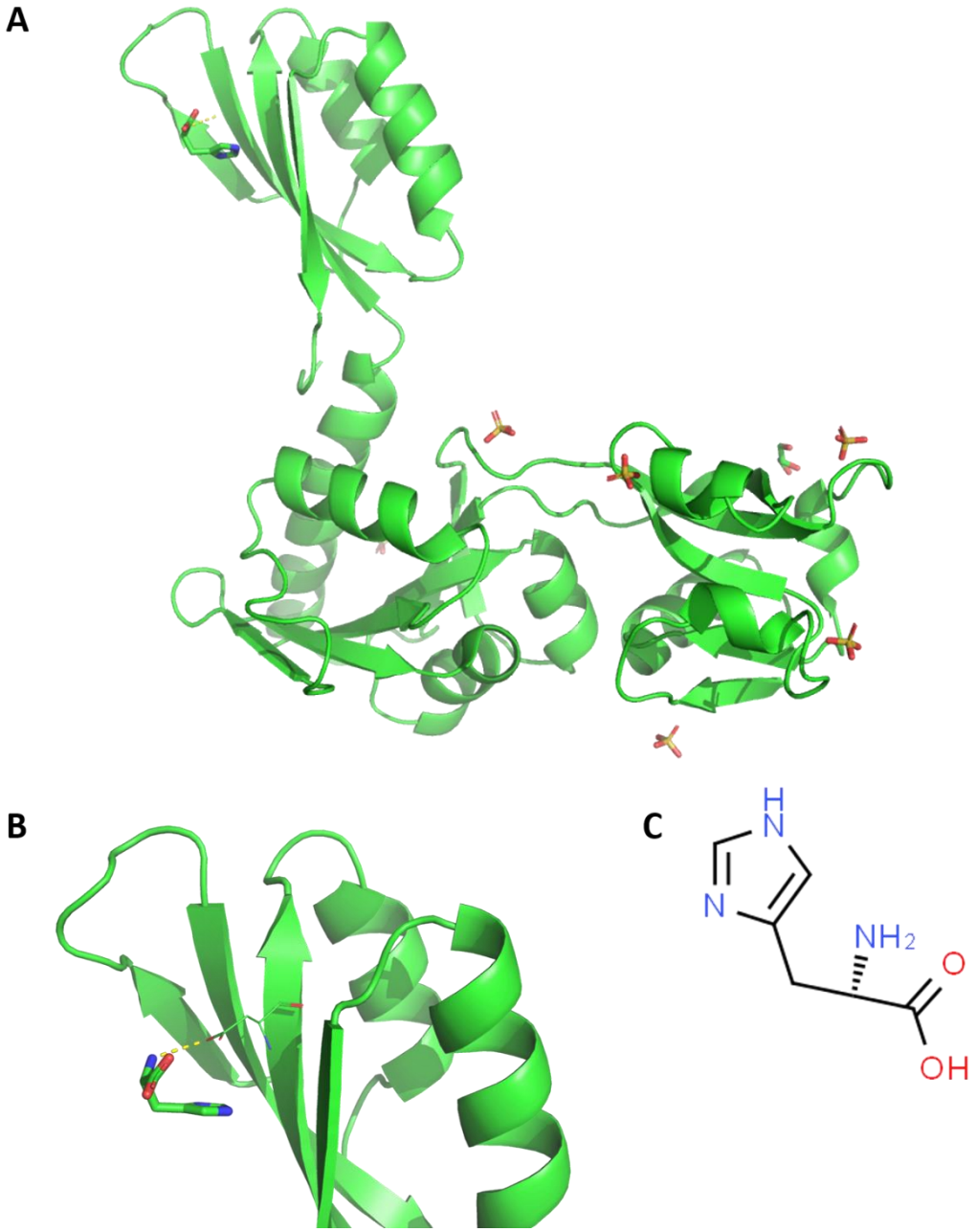

Figure 1.10. Crystal structure of *mtu*ATP-PRT (PDB: 5LHU). A: Single chain of *mtu*ATP-PRT with L-His bound to the βαββαβ fold of the regulatory domain III. B: close inspection of L-His interacting with Asp218 of domain III. C: 2D chemical structure of L-His.

1.3.2.2 Allosteric activation by 3-(2-Theinyl-L-Alanine)

Using Compound Screening in the Presence of an Inhibitor (CoSPI) to screen L-His analogues, Pisco *et al.* (2017) identified TIH as a non-essential activator of

*mtu*ATP-PRT. L-His would be considered an essential inhibitor of ATP-PRT due to its role in the biochemical pathway ATP-PRT is associated with. TIH is designated as a non-essential activator because it can bind and alter ATP-PRT activity without being essential for normal enzyme activity. Using steady-state kinetics, TIH was shown to increase activity of *mtu*ATP-PRT by 490 %, significantly increasing the turnover of PR-ATP. TIH also functions as a rescue molecule in the presence of L-His, restoring the L-His bound ATP-PRT form from 0 % activity to 100 %. This rescuing effect of TIH seems to occur when L-His and TIH are simultaneously bound into ATP-PRT multimer (Pisco *et al.*, 2017).

In terms of allosteric mechanism, TIH binds to the C-terminal regulatory domain of *mtu*ATP-PRT just as L-His does (**Figure 1.11**) and stimulates a twist to the regulatory domain. Both L-His (**Figure 1.10**) and TIH (**Figure 1.12**) show interactions to allosteric ligand *via* Asp218 of the regulatory domain. In L-His inhibited *cje*ATP-PRT this twist creates a steric clash at the active site therefore disallowing simultaneous binding of substrates. This steric clash would also be observed in allosteric activation by TIH in *mtu*ATP-PRT based on crystal structures, yet catalysis is increased. This apparent contradiction in the structural changes associated with activator and inhibitor binding, in the available crystallographic data suggests gaps in our understanding of allosteric signalling in ATP-PRT.

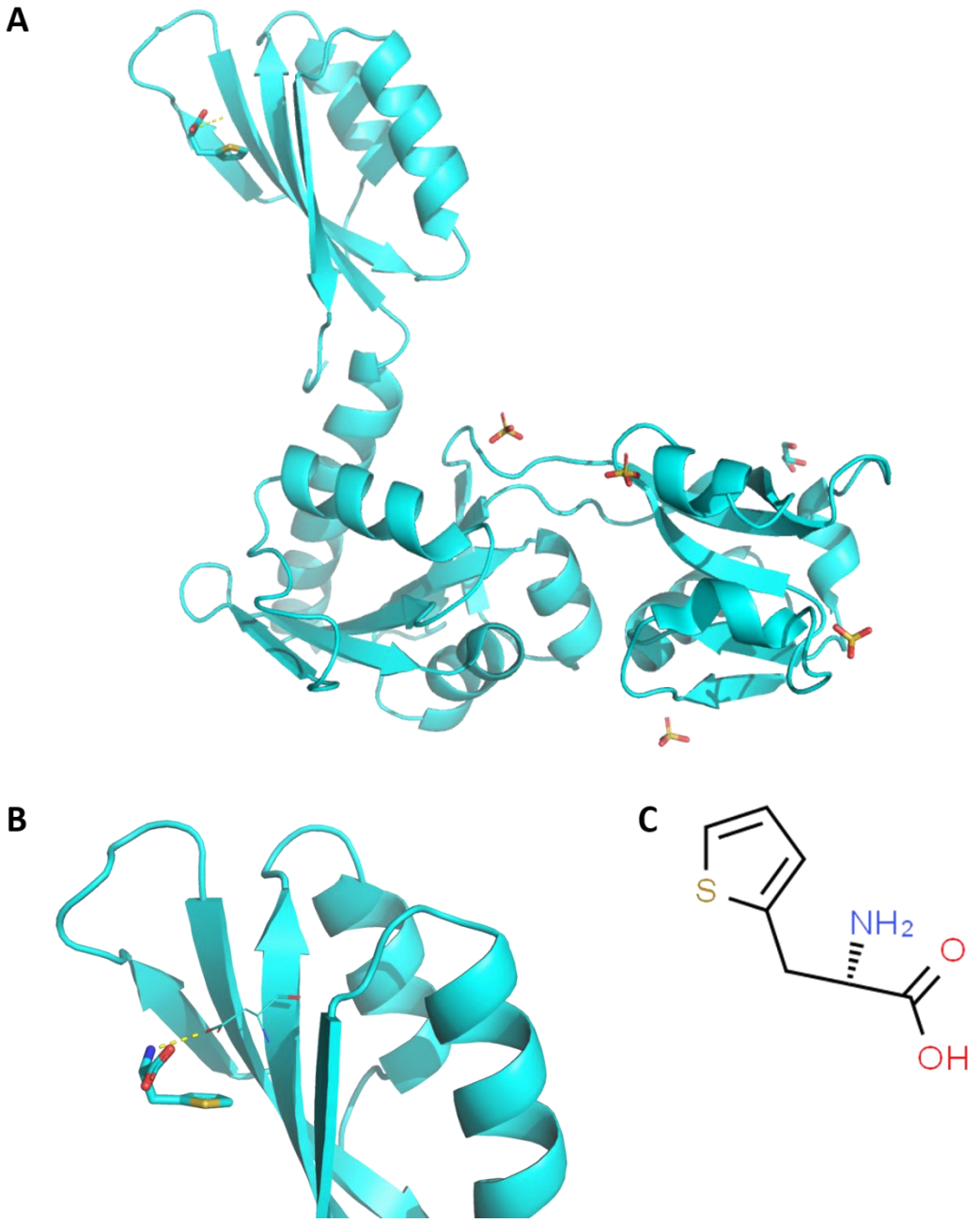

Figure 1.11. Crystal structure of *mtu*ATP-PRT (PDB: 5LHT). A: Single chain of *mtu*ATP-PRT with 3-(2-Thienyl)-L-alanine (TIH) bound to the βαββαβ fold of the regulatory domain III. B: close inspection of TIH interacting with Asp218 of domain III. C: 2D structure of TIH, note structural similarity to L-Histidine (**Figure 1.10**).

1.3.2.3 Allosteric regulation in ATP-PRT is not explained by conformational allosteric models

Based on X-ray crystallography, both L-His and TIH binding elicit a superimposable tensed conformation to *mtu*ATP-PRT. TIH also binds in the same

allosteric site to L-His, and both allosteric ligands exhibit the same residue interactions (Pisco *et al.*, 2017). A single residue difference has been identified between *mtu*ATP-PRT structures (PDB: 5LHU and 5LHT) with Asp216 orientating towards the thiophenyl ring of TIH in the allosterically active crystal structure. However, on mutating this residue only allosteric inhibition was affected, and allosteric activation remained.

The similarity of both crystal structures is surprising, considering their opposing effects on catalytic rate. With both structures adapting a **T** state this opens the possibility of ATP-PRT allosteric regulation being a dynamic process that does not fit into traditional allosteric models and thus requires a detailed investigation to assess structural, kinetic, and dynamic changes that occur between allosteric states.

1.4 Temperature dependence of enzymes

Enzymes catalyse many crucial biological functions and have the capacity to increase reaction rates up to 10^{17} times faster than uncatalysed reactions (Fersht, 1999; Kamerlin & Warshel, 2010). Considering the catalytic power enzymes have, the phenomena that contribute to this catalysis should be deeply understood. Chemical and enzymatic reactions occur at different rates and are strongly influenced by temperature. In enzyme catalysed reactions, this temperature relationship results in enzyme rates peaking at an optimal temperature (T_{opt}) and declining thereafter. The effect of temperature on enzymatic and chemical rates has been investigated for years and is classically explained by enzyme denaturation above specific temperatures. However, curvature in the enzyme rate-temperature plot exists at temperatures below denaturation temperatures and thus requires an explanation that does not invoke denaturation. Macromolecular Rate Theory (MMRT) provides an extension to classical models of chemical reaction rates and has been applied to the temperature dependence of enzymatic rates (Hobbs *et al.*, 2013; Arcus & Pudney, 2015; Arcus *et al.*, 2016; Van Der Kamp *et al.*, 2018).

1.4.1 Temperature dependence of enzyme reactions

The Arrhenius Equation (**Equation 1.1**) was first formalised by Svante Arrhenius in the late 19th century to illustrate the relationship between chemical reaction rates and temperature. The Arrhenius equation (**Equation 1.1**) describes the increase in reaction rate (*k*) with temperature (T) in the context of an activation energy barrier where A is a pre-exponential factor, E_A the activation energy, and R the universal gas constant.

Equation 1.1 Arrhenius Equation.

$$\boldsymbol{k = Ae^{-\frac{E_A}{RT}}}$$

The Arrhenius equation implies that chemical reaction rates increase exponentially with temperature (**Figure 1.12**) as the activation energy (E_A) in the reaction is always positive (Fersht, 1999). Increasing temperatures provide increasing energy for a reaction to reach the transition state and thus increases rates (Justi & Gilbert, 1999). Increased temperature increases the kinetic potential of reactants and therefore increases the probability of reactants colliding with enough energy to surpass the activation energy barrier.

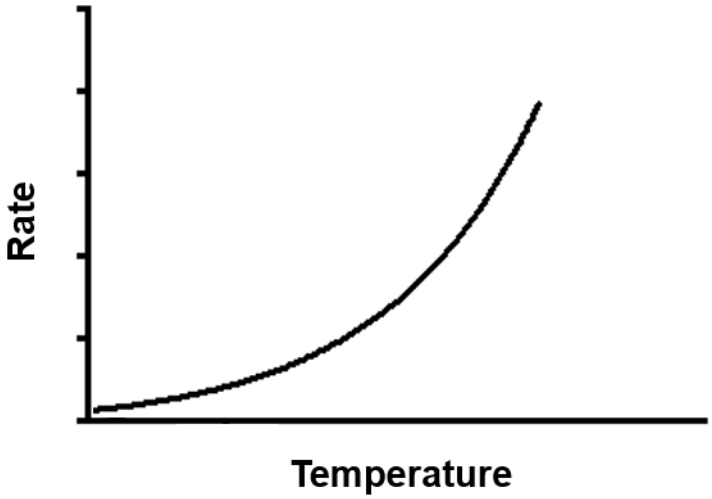

Figure 1.12. Arrhenius function showing positive and exponential increase of non-catalysed reaction rates in response to increased temperature.

Eyring and Polanyi (Eyring, 1931) then expanded on the function of activation energy from the Arrhenius equation, defining the change to Gibb's free energy between the reactants and the transition state (TS) (**Equation 1.2**). The pre-exponential factor was expanded to include κ as the transmission co-efficient, k_B as Boltzmann's constant, and temperature over Planck's constant (h) (Fersht, 1999). Expanding on the original Arrhenius equation, the Eyring-Polanyi equation provided an improved description of reaction rates in simple chemical reactions by defining the transition state energy barrier ($\Delta G^{\ddagger}$). Both the Arrhenius and Eyring equations are traditionally applied to chemical systems and simple biological reactions, with great descriptive power.

Equation 1.2 Eyring-Polanyi equation. k is the rate constant, κ is transmission co-efficient, k_B is Boltzmann's constant, h is Planck's constant, $\Delta G^{\ddagger}$ is Gibb's free energy between the reactants and TS, R is the universal gas constant, and T is temperature (Fersht, 1999).

$$\boldsymbol{k = \left(\frac{\kappa k_B T}{h}\right) e^{-\Delta G^{\ddagger}/RT}}$$

An enzyme following Arrhenius or Eyring models would be expected to show exponential increases to rate in response to increasing temperature (**Figure 1.12**), however many enzymes do not show typical Arrhenius behaviour and instead show non-exponential rate increases (**Figure 1.13**). In addition enzyme reactions across increasing temperature ranges exhibit a T_{opt} with subsequent decreased reaction rates. The clear difference in chemical *versus* enzyme catalysed reaction rates suggests enzyme catalysed reactions have a more complicated temperature dependence than previously described by Arrhenius and Eyring equations (Arcus *et al.*, 2016).

Curvature of enzymatic rates after T_{opt} could be attributed to global denaturation of the enzyme, and this is the typical textbook description (**Figure 1.13**). However, protein denaturing temperatures do not always coincide with the temperature where rates start to decrease, as would be a requirement for this denaturation dependent explanation.

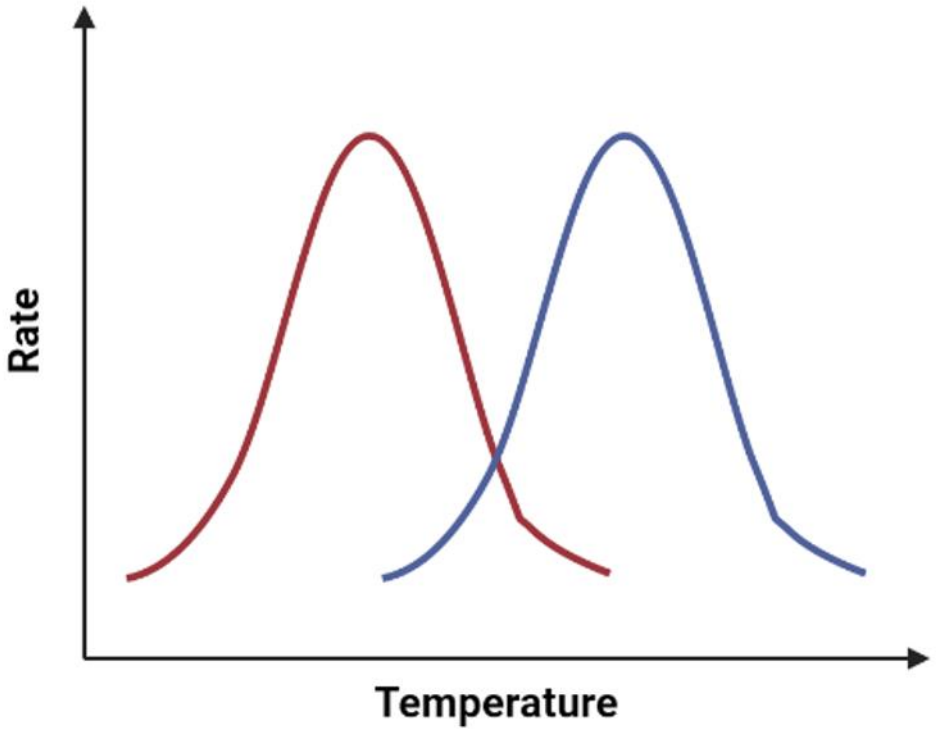

Figure 1.13. Rate against temperature for a psychrophilic (red) and mesophilic (blue) enzymes.

For example, psychrophilic α-amylase has a T_{opt} of 25-30 °C and decreasing curvature in rates is seen thereafter. Whereas the melting temperature (T_m) of psychrophilic α-amylase is 44 °C, far beyond the onset of decreasing rates (D'Amico *et al.*, 2006; Sočan *et al.*, 2020). This T_{opt} - T_m disparity is not enigmatic, and has been observed in mesophilic, psychrophilic, and extremophilic enzymes which demonstrates a T_{opt} below the observed T_m indicating curvature is not generally a function of enzyme unfolding (Feller & Gerday, 2003; D'Amico *et al.*, 2006; Hobbs *et al.*, 2013).

Daniel *et al.* (2001) introduced the equilibrium model (T_{eq}) to describe a model of enzyme behaviour with temperature. T_{eq} provides a model for enzyme curvature without denaturation. This model postulates an enzyme to be in an equilibrium between folded active and folded but inactive, enzymatic forms (Daniel & Danson, 2010). In the T_{eq} model, temperature drives the equilibrium towards either active or inactive states, and declining rates after T_{opt} are attributed to an equilibrium shift towards the inactive state (Daniel *et al.*, 2001).

1.4.1.1 Macromolecular Rate Theory (MMRT)

MMRT describes typical curvature of enzymatic rates across temperature in the absence of enzyme denaturation. Expanding on the traditional Eyring equation (**Equation 1.3**), with the inclusion of an activation heat capacity ($\Delta C_p^{\ddagger}$), MMRT offers a novel equation to describe enzyme behaviour over large temperature ranges.

Equation 1.3 Expanded Eyring Equation.

$$\boldsymbol{ln(k) = ln\left(\frac{\kappa k_B T}{h}\right) - \frac{\Delta H^{\ddagger}}{RT} + \frac{\Delta S^{\ddagger}}{R}}$$

Equation 1.4 Temperature dependence of H.

$$\boldsymbol{\Delta H_T^{\ddagger} = \Delta H_{T_0}^{\ddagger} + \Delta C_p^{\ddagger}(T - T_0)}$$

Equation 1.5 Temperature dependence of S.

$$\boldsymbol{\Delta S_T^{\ddagger} = \Delta S_{T_0}^{\ddagger} + \Delta C_p^{\ddagger} ln\left(\frac{T}{T_0}\right)}$$

Equation 1.4 and Equation 1.5 show the expansion of H and S to be temperature dependent based on the introduced heat capacity term. The temperature dependence of $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ are both defined by $\Delta C_p^{\ddagger}$ (change in heat capacity between reactant and TS). These expanded functions (**Equation 1.3**, **Equation 1.4**, **Equation 1.5**) form the first-generation model of MMRT (**Equation 1.6**, MMRT 1.0, Temperature independent model).

There are two published models of MMRT; the first generation model has $\Delta C_p^{\ddagger}$ independent of temperature (**Equation 1.6**) (Hobbs *et al.*, 2013) and the second generation model has a temperature dependent $\Delta C_p^{\ddagger}$ term (Hamill, 2020) (**Equation 1.7**, MMRT 1.5). Each MMRT model includes elements of thermodynamic properties, and can extrapolate data on activation heat capacity, entropy and enthalpy that can contribute to an understanding of underlying thermodynamic drivers to curvature in temperature-rate plots.

Equation 1.6 Temperature independent MMRT (MMRT 1.0) equation from (Hobbs *et al.*, 2013). Where k = rate; k_B = Boltzmann constant; h = Planck's constant; $\Delta H^{\ddagger}_{T0}$ = enthalpy change at a reference temperature T_0; R = ideal gas constant; $\Delta S^{\ddagger}_{T0}$ = entropy change at T_0.

$$ln(k) = ln\left(\frac{k_B T}{h}\right) - \frac{\Delta H^{\ddagger}_{T_0} + \Delta C^{\ddagger}_p (T - T_0)}{RT} + \frac{\Delta S^{\ddagger}_{T_0} + \Delta C^{\ddagger}_p ln\left(\frac{T}{T_0}\right)}{R}$$

Equation 1.7 Temperature dependent MMRT (MMRT 1.5) equation from (Hamill, 2020). Where k = rate; k_B = Boltzmann constant; h = Planck's constant; $\Delta H^{\ddagger}_{T0}$ = enthalpy change at T_0; R = ideal gas constant; $\Delta S^{\ddagger}_{T0}$ = entropy change at T_0. The $\Delta C_p^{\ddagger}$ term is now linearly dependent with temperature.

$$ln(k) = ln\left(\frac{\kappa k_B T}{h}\right) - \frac{\Delta H^{\ddagger}_{T_0} + \left[\Delta C^{\ddagger}_{p,T_0} + A(T - T_0)\right](T - T_0)}{RT} + \frac{\Delta S^{\ddagger}_{T_0} + \left[\Delta C^{\ddagger}_{p,T_0} + A(T - T_0)\right] ln\left(\frac{T}{T_0}\right)}{R}$$

MMRT has been used to explain thermodynamic properties between enzymes adapted to different temperatures. Arcus and Mulholland (2020) described an enthalpy–entropy trade-off in the context of psychrophilic, mesophilic, and thermophilic enzymes. Lowering the enthalpic barrier of a reaction results in a lower free energy barrier at low temperatures. However this decrease in $\Delta H^{\ddagger}$ is generally traded at the expense of an increase in $\Delta S^{\ddagger}$ which is matched with a more negative $\Delta C_p^{\ddagger}$ (Arcus & Mulholland, 2020). Thermophilic enzymes are suggested to be more rigid, like an allosterically inhibited state, and psychrophilic enzymes more flexible. Therefore, this enthalpy-entropy trade-off can be used to explain rate changes between psychrophilic, mesophilic, and thermophilic reaction rates and similarly, may be used to explain changes to catalytic rate of allosterically inhibited, APO, and allosterically activated enzyme catalysed reactions (**Figure 1.14**).

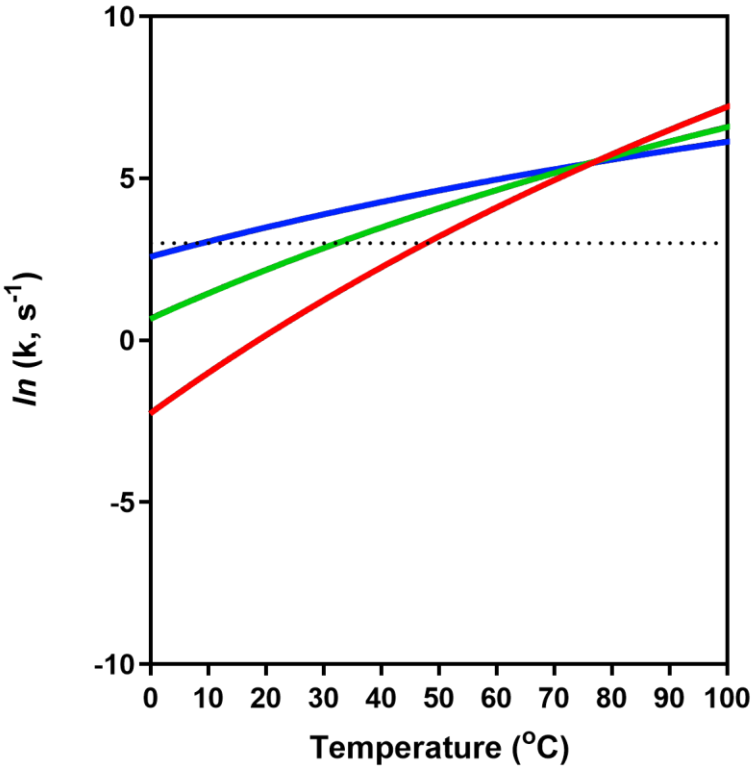

Figure 1.14. Theorised temperature dependence of an allosteric enzyme adapted from (Arcus & Mulholland, 2020) showing hypothetical reaction rates of an allosteric enzyme across temperature according to the Eyring equation. Enzyme rates for the same reaction are illustrated for hypothetical APO protein (green), allosterically activated (blue), and allosterically inhibited (red) protein. The decreasing slope for each allosteric state illustrates the enthalpy-entropy trade off, in each allosteric state. The dotted line indicates a k_{cat} of 20 s^{-1} ($\ln k_{cat}$ =3.0). The rates coincide at ~80 °C.

1.4.2 MMRT as a tool to better understand allostery

MMRT has been used to link the temperature-dependent curvature in individual enzymes of a metabolic pathway to the curvature of net pathway rates, and to the temperature dependence of organism growth rates (Schipper *et al.*, 2014; Prentice *et al.*, 2020). However, the enzymes were all studied with simple enzyme systems (Hobbs *et al.*, 2013) or on a global scale (Prentice *et al.*, 2020) without taking into account any regulatory processes such as allostery. Application of MMRT to a large enzyme with numerous regulatory mechanisms would extend this analysis to better understand the full complexity of enzymes in metabolism.

Exploration of the temperature dependence of a large enzyme system with allosteric regulation would allow the exploration of changes to activation entropy, enthalpy,

and heat capacity in different ligand bound states. This would potentially provide insight into the dynamic mechanisms driving its activity and regulation in a manner not achieved using traditional steady-state kinetic analysis.

1.4.3 Mycobacterium tuberculosis, drug resistance and ATP-PRT

M. tuberculosis was first identified in 1892 by Robert Koch, and is a rod shaped, pathogenic bacterium with a lipid rich, waxy cell wall structure. It is the causative agent of tuberculosis (TB) (Cambau & Drancourt, 2014). TB is a widespread respiratory pathogen with 9 million individuals thought to contract TB annually. The lipid rich wall of *M. tuberculosis* contributes to its pathogenicity and necessitates antibiotic treatments over long periods, with isoniazid, rifamycin, pyrazinamide, and ethambutol administered over a time period of six to nine months (Koul *et al.*, 2011). This complex and drawn-out treatment regime results in non-adherence by patients and contributes to drug resistance of *M. tuberculosis*. Due to this drug resistance, there are now multidrug resistant (MDR) and extensively drug resistant (XDR) strains of TB which do not respond to isoniazid and rifampicin, and second line treatments including fluroquinolone, kanamycin, amikacin, and capreomycin in the case of XDR (Gandhi *et al.*, 2010).

The origin of MDR and XDR TB has accelerated the need for new pharmacological approaches to the treatment of TB. ATP-PRT is touted as a drug target in *M. tuberculosis* due to its critical role in the essential histidine biosynthesis pathway. Absent in humans, ATP-PRT is an attractive drug target due to its many regulatory mechanisms, essentiality, and possibility to avoid side effects on to the host. However, before drug development, it is essential to understand the mechanisms behind each mode of regulation.

Gaining a deep understanding of allosteric mechanism has both theoretical value, and potentially value in the area of drug discovery – including treatment of *M. tuberculosis*. Allostery is a direct and efficient method of biological regulation and is prevalent in many major metabolic functions and signal transduction pathways, therefore is increasingly becoming exploited for therapeutics. Existing allosteric drugs work as inhibitors, activators, or neutral modulators and function to

alter the regular biological function of an enzyme or regulator. Pharmacologically, allosteric modulating drugs have higher binding affinities, greater selectivity, and offer lower side effects and toxicity compared to drugs that target active sites (Lu *et al.*, 2014). With increasing MDR and XDR, *M. tuberculosis* strains allosteric drug discovery could offer a solution. *mtu*ATP-PRT, absent in humans could be exploited for drug development, offering a specific binding target with low toxicity to the host.

1.5 Research objectives

This research investigates the origins and regulation of ATP-PRT. Phylogenetic analysis will be used to explore the origins and evolution of ATP-PRT isoforms and their respective regulatory domains. It also aims to apply a novel approach (MMRT) to gain a deeper understanding of enzyme mechanism, biological rates, and dynamic allosteric regulation in ATP-PRT. *mtu*ATP-PRT is used as both a model of a more complex enzyme due to its multimeric nature and complex catalytic reaction, but also to probe dynamic origins of allosteric inhibition and activation present in ATP-PRT, and ultimately allosteric regulation mechanisms in general.

1. Use phylogenetic approaches to investigate the phylogeny of ATP-PRT isoforms across bacterial lineages;

2. Test the scalability of MMRT on ATP-PRT as a complex enzyme model;

3. Use MMRT as a tool to test our understanding allosteric mechanisms;

4. Investigate the mechanism of allosteric signal transduction in ATP-PRT *via* the generation of allosteric resistant mutants.

Chapter 2
Evolution of ATP-PRT

This chapter explores the possible evolutionary pathways of the ATP-PRT associated genes; *hisG$_L$*, *hisG$_S$*, and *hisZ*, using a phylogenetic analysis of the two ATP-PRT isoforms. The aim is to elucidate the ancestral form(s) of ATP-PRT and gain insight into the possible evolutionary advantages of each isoform.

2.1 Introduction to ATP-PRT

Two forms of ATP-PRT are reported in literature, a long form and a short form, both functioning as pentosyltransferases. Both ATP-PRT forms catalyse the reversible condensation of ATP and PRPP in the *de novo* synthesis of histidine. Initiating a metabolic pathway that uses 41 ATP equivalents (Brenner & Ames, 1971) requires substantial regulation and as a result both ATP-PRT isoforms are subject to allosteric regulation by the end product histidine (Mittelstädt *et al.*, 2018). Both ATP-PRT isoforms share a two-domain active site, and are catalytically active with the active site alone, however, the isoforms diverge in their regulatory domains (**Figure 1.8** and **Figure 2.1**).

Longform ATP-PRT (ATP-PRT$_L$) is encoded by *hisG$_L$* and makes up 46 % of all ATP-PRTs (based on PFAM database entries). *hisG$_L$* encodes a 280-310 amino acid protein consisting of domain I and domain II, forming the bi-lobal catalytic core. The regulatory domain III comes from an **a**spartate kinase, **c**horismate mutase and **T**yrA (ACT) domain extension at the c-terminus of *hisG$_L$* (denoted HisG_C). HisG_C forms a well-defined allosteric binding pocket for allosteric feedback inhibition by L-His (Champagne *et al.*, 2006; Mittelstädt *et al.*, 2016). The catalytically active ATP-PRT$_L$ multimer is a hexamer, containing six catalytic subunits and six allosteric binding sites (**Figure 1.8**).

Short-form ATP-PRT (ATP-PRT$_S$) is comprised of two gene products (Mittelstädt *et al.*, 2018); *hisG$_S$* and *hisZ*. *hisG$_S$* is a truncated *hisGL*, approximately 200 residues

long, and entirely lacks the HisG_C domain and its allosteric binding site (Weiner, 1999; Mittelstädt *et al.*, 2016). *hisG*$_S$ encodes a very similar bi-lobal catalytic site, however, the multimeric structure only contains four catalytic subunits (**Figure 1.8**). HisG$_S$ forms a hetero-octameric complex with four tRNA synthetase like *hisZ* molecules to form the active quaternary structure, providing allosteric regulation and enhanced catalysis (Livingstone *et al.*, 2016; Mittelstädt *et al.*, 2018).

2.1.1 Structure of ATP-PRT isoforms

Structurally, in ATP-PRT$_L$ the active site domains are in the centre of the hexamer which assembles as a dimer-of-trimers. This organisation buries the active site residues inside the hexamer. However, in ATP-PRT$_S$, the active sites are orientated at the surface of the complex and are relatively solvent exposed, as the multimeric structure is formed through the interactions with *hisZ*. ATP-PRT$_S$ orientate as a catalytic dimer flanking an X-shaped core of *hisZ* regulatory subunits (**Figure 2.1**) (Champagne *et al.*, 2005; Vega *et al.*, 2005).

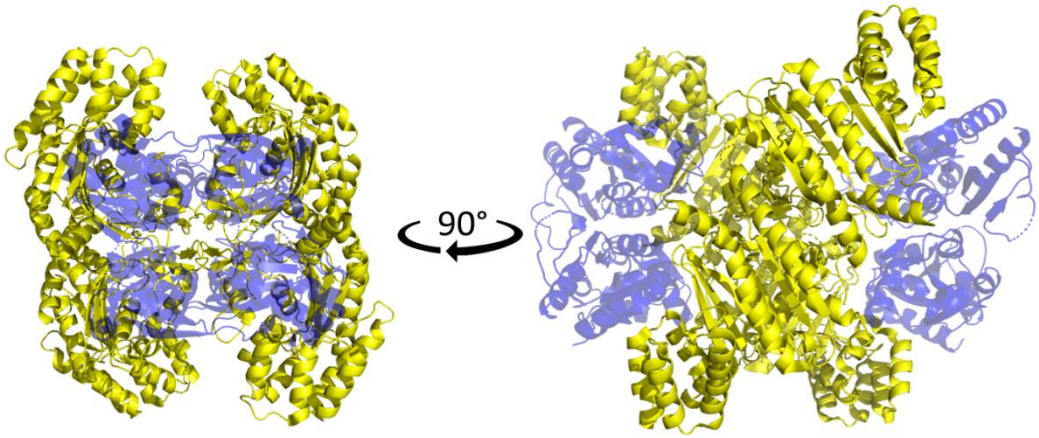

Figure 2.1 Subunit orientation of ATP-PRT$_S$ from *L.lactis* (PDB: 1Z7M). The semi-transparent catalytic core (blue ribbon) arranges in two dimer pairs flanking the tetrameric interface of *hisZ* (yellow ribbon).

hisG$_L$ and *hisG*$_S$ share approximately 25 % sequence identity for their catalytic subunits (Champagne et al., 2006), however their regulatory domains HisG_C and *hisZ* are not homologous. *hisZ* is instead homologous to the catalytic domain of Histidyl-tRNA synthetases (*hisS*), sharing key residues for the binding of histidine (Weiner, 1999). Unusually, when investigating the binding of histidine in *hisZ* it

has been shown that these shared residues do not interact with histidine in the *hisG$_S$/hisZ* complex (Vega et al., 2005).

The gene fragment encoding the C-terminal regulatory domain (HisG_C) shares structural and sequence homology to ACT domains (Lang *et al.*, 2014; Mittelstädt *et al.*, 2018). The ACT protein domain is commonly found in metabolic proteins, and specifically linked to enzymes regulated by amino acid concentration (Chipman & Shaanan, 2001). The sequence identity between ACT domains is low, likely due to the age and variety of functions ACT domains serve. Regardless of sequence identity, the structural similarities between ACT-like domains with a βαββαβ topology is seen in the regulatory C-terminus of ATP-PRT$_L$ structures and other ACT-like domains (Lang *et al.*, 2014). ACT domains can be artificially fused to proteins with no existing regulatory domain, and are able to confer allosteric regulation (Cross *et al.*, 2013) thus demonstrating the ease with which this domain can be acquired for regulation through the course of evolution.

2.1.2 Existing evolutionary models

The evolutionary theories regarding ATP-PRT isoforms have changed significantly over time as more information has been gathered surrounding the function and regulation for both *hisZ* and *hisG$_L$*. Previously *hisG$_L$* was thought to belong to just Archaea, Eukarya, and Eubacteria (Weiner, 1999), however, there is no clear distinction between the existence of short- and long-forms in any phylogenetic class. HisG$_L$ and HisG$_S$ are distributed across almost all archaeal and proteo-bacterial groups; with HisG$_L$ dominating proteo-bacterial ATP-PRT, and *hisG$_S$* dominating within the Archaea and Eukarya domains. There is also a select group of organisms including the Geobacter clade, *Candidatus Accumulibacter phosphatis*, and *Desulfotomaculum acetoxidan*s that contain functional forms of both *hisG$_L$*, and the *hisG$_S$*/*hisZ* genes (Aklujkar, 2011). These have not been included in any previous published phylogenetic trees of ATP-PRT.

Bond and Francklyn (2000) and Vega *et al.* (2005) both suggest short-form ATP-PRT as the ancestral form with *hisZ* gene loss occurring over time. Bioinformatic analysis suggests a convergent evolution of regulatory systems in these two proteins; implicating *hisG$_S$* as the phylogenetically older form (Vega *et*

al., 2005). The two forms likely acquired both regulatory allostery, and enhanced catalysis by the recruitment of an ACT domain and a histidyl-tRNA synthetase-like subunit in two evolutionary events to form the two diverse groups we observe today. These ancestral forms have since diverged, creating both the *hisZ* monophyletic group separate from *hisS* (Bond & Francklyn, 2000), and the low sequence similarity in ACT domains (Mittelstädt *et al.*, 2018).

The simple explanation for both isoforms is to assume either short or long ancestors that then diverged with the adoption of separate regulatory domains. However, there is now evidence for a wide range of variants (*e.g.*, some organisms containing both forms, multifunctional fused forms, ATP-PRT's lacking regulatory domains entirely, and loss of *hisG* genes in their entirety). This complexity suggests a more complex evolutionary story with multiple events.

2.1.3 L-His inhibition of ATP-PRT forms

In assessing the evolutionary gain of either isoform, understanding any trends in their catalysis and regulation is critical. Kinetic information on L-His inhibition for *Thermotoga maritima, Lactoccocus lactis, Salmonella typhimurium, Camplyobacter jejuni, Aridopsis thaliana, Psychrobacter arcticus,* and *Mycobacterium tuberculosis* was collated from the literature to assess any differences in histidine sensitivity across different isoforms of ATP-PRT (Table **2.1**).

Table 2.1. Inhibition by L-His data of different ATP-PRT isoforms as reported in literature.

Organism	L-His (μM)	Method of inhibition	Isoform	Reference
Mycobacterium tuberculosis	36	K_i (PRPP)	Long	(Pisco *et al.*, 2017)
Campylobacter jejuni	40-44	K_i (PRPP)	Long	(Mittelstädt *et al.*, 2016)
Aridopsis thaliana (ATP-PRT1)	45	IC_{50}	Long	(Ohta *et al.*, 2000)
Salmonella typhimurium	70-100	K_i	Long	(Martin, 1963; Kleeman & Parsons, 1976;

				Morton & Parsons, 1976)
Aridopsis thaliana (ATP-PRT2)	320	IC_{50}	Long	(Ohta *et al.*, 2000)
Psychrobacter arcticus $HisG_S$	0.25 k_{cat} (s^{-1})*	Steady state kinetics (0 µM L-His)	Short	(Fisher *et al.*, 2018)
Psychrobacter arcticus $HisG_S$/HisZ complex	1.5 k_{cat} (s^{-1})*	Steady state kinetics (0 µM L-His)	Short	(Fisher *et al.*, 2018)
Lactococcus lactis	81.1	K_i	Short	(Champagne *et al.*, 2006)
Thermotoga maritima	350	K_i	Short	(Vega *et al.*, 2005)

* No available L-His data available for *P. arcticus*. k_{cat} (s^{-1}) provided for *P. arcticus* with and without *hisZ* (regulatory domain) complex to indicate the increased catalytic effect of *hisZ* on ATP-PRT_S.

Long and short isoforms exhibit broad sensitivities to histidine ranging from 30-350 µM, substantially lower than the intracellular L-His concentration of 15 – 100 µM in *S. typhimurium* and *E. coli* (Winkler & Ramos-Montañez, 2009). However, it would be inappropriate to draw conclusions from these data due to variations in data collection (pH, temperature, and buffer). There is no K_i presented for *P. arcticus* but there is steady-state kinetics detailing the enhanced catalysis of $HisG_S$ offered by *hisZ*. Overall there is no clear kinetic data to suggest differences in regulation that may offer different selection advantages across species. However it is difficult to compare these due to the amount of available data and variation in collection conditions.

2.1.4 Aim of chapter

Using phylogenetic approaches, this chapter investigates the evolutionary origins of the two ATP-PRT isoforms. The aim of this chapter is to determine if there is any evidence to suggest a single isoform as the ancestor; if (and what) environmental pressures may have resulted in the divergence of the second isoform, and what biological pressures persist today that continue the divergence of these two ATP-PRT isoforms in bacteria.

2.2 Methods and Results

Different phylogenetic methods were undertaken in the process of generating phylogenetic trees to investigate ATP-PRT evolution and distribution of isoforms across bacteria. Each phylogenetic method was trialled with subsets of the species across a subset of the tree of life, and the final trees presented here were generated using methods that were; sufficient, fast, and produced trees with appropriate topology. Methods and results have been merged here as this phylogenetic analysis used a holistic approach, where the results from one, led to the method of another.

2.2.1 Guided tree

Due to the abundance of ATP-PRT's across archaea, bacteria, and plants, the retrieval of sequences was cumbersome. With the variety of ATP-PRT isoforms across each domain, and the sheer number of bacterial sequences available, bacterial species were limited to the representative prokaryotic tree taken from Battistuzzi *et al.* (2004). The species from this prokaryotic tree represent a broad phylogenetic distribution and are well studied organisms, with well annotated genomes. This included a total of 54 sequences, distributed across various bacterial groups. *hisG* sequences of each bacterial species were annotated as either $hisG_L$, $hisG_S$, (or absent) based on sequences retrieved from InterPro and UniProt (**Figure 2.2**).

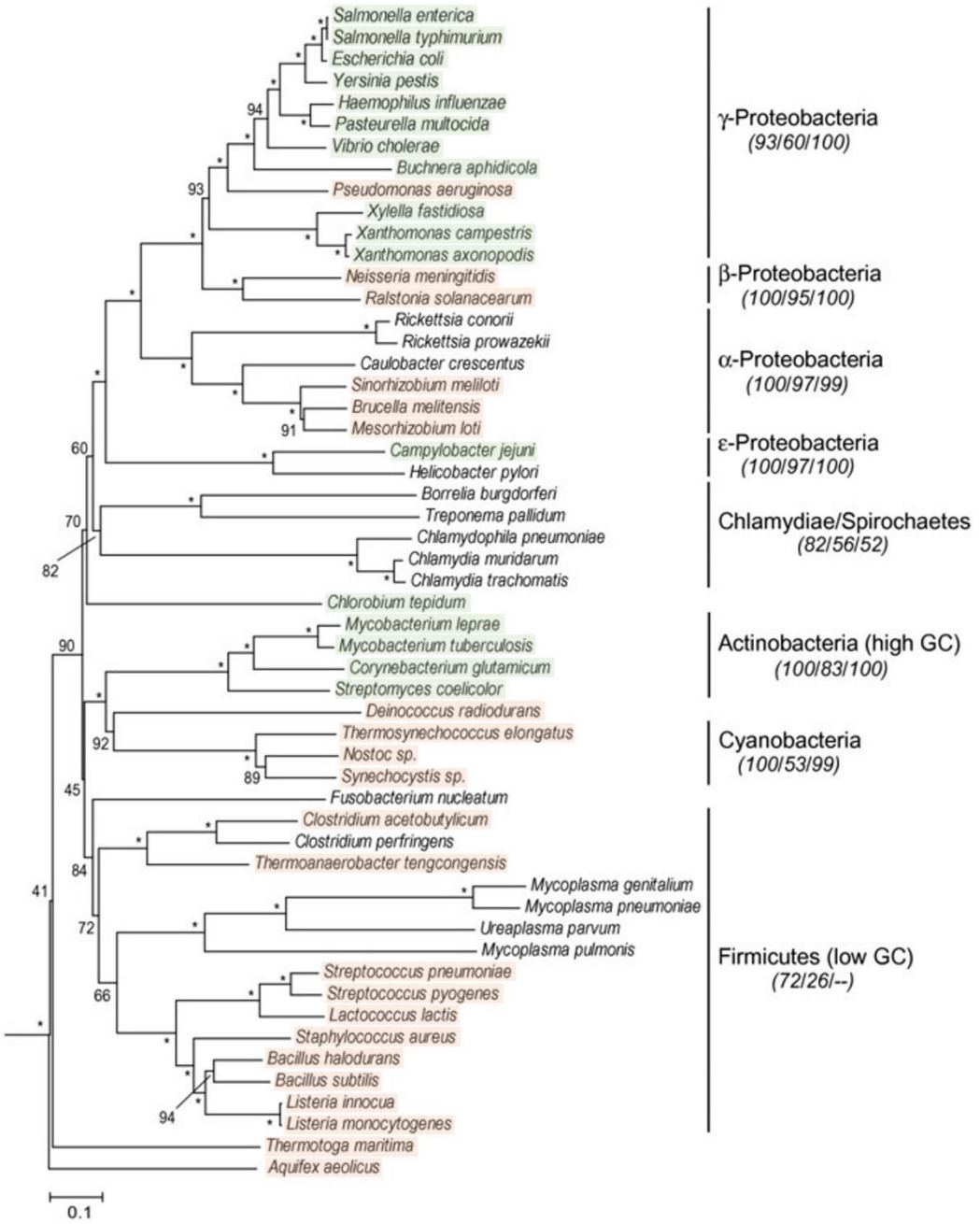

Figure 2.2 Prokaryotic tree taken from (Battistuzzi *et al.*, 2004) annotated with *hisG* gene type. $hisG_L$ (green), $hisG_S$ (orange), organisms with no *hisG* genes present are uncoloured. Species were defined as $hisG_L$ based on a sequence length of 200+ residues and the absence of *hisZ*. $hisG_S$ annotations were defined based on the presence of a truncated *hisG* gene (<200 residues) and the presence of *hisZ*. Annotations were based on sequence annotations retrieved from InterPro and UniProt.

A total of 57 organisms (**Table 2.2**) were selected for further investigation and tree building. Selection of species was based on those present in **Figure 2.2**, plus organisms with solved crystal structures (*Psychrobacter arcticus*). *Eggerthella lenta, Collinsella aerofaciens Deinococcus proteolyticus,*

Geobacter sulfurreducens, and Slackia heliotrinireductens were added due to evidence of containing either; a multifunctional fusion protein (fused *hisG*/*hisZ* genes) or the presence of both ATP-PRT isoforms.

Table 2.2 Species collected from Prokaryotic tree (Battistuzzi et al., 2004) and species selected from known structures annotated for presence of $hisG_L$, $hisG_S$, and *hisS*. Organisms that did not contain either form of HisG on UniProt were donated as 'Absent' on the basis of the absence of a critical step in the only known His biosynthesis pathway. Definition of isoform was made based on the genes annotated from InterPro and UniProt

Species	$hisG_L$	$hisG_S$	*hisZ*	**Notes**
Aquifex aeolicus		+	+	
Bacillus subtilis		+	+	
Borrelia burgdorferi				Absent
Brucella melitensis		+	+	
Buchnera aphidicola	+			
Campylobacter jejuni	+			
Caulobacter crescentus	+			
Chlamydia trachomatis				Absent
Chlamydophila pneumoniae				Absent
Chlorobium tepidum	+			
Clostridium acetobutylicum		+	+	
Clostridium perfringens				
Collinsella aerofaciens		+	+	Multifunctional fusion protein
Corynebacterium glutamicum	+			
Deinococcus proteolyticus	+			
Deinococcus radiodurans		+	+	
Desulfofarcimen acetoxidans	+	+	+	Dual forms
Eggerthella lenta		+	+	Multifunctional fusion protein
Escherichia coli	+			
Flavonifractor plautii		+		
Geobacter sulfurreducens	+	+	+	Dual forms
Haemophilus influenzae	+			
Helicobacter hepaticus		+		
Helicobacter pylori				Absent
Lactococcus lactis		+	+	
Leptospira interrogans		+	+	
Listeria monocytogenes		+	+	
Mesorhizobium loti (japonicum)		+	+	
Mycobacterium tuberculosis	+			
Mycoplasma pneumoniae				Absent
Neisseria gonnorhea		+	+	
Nostoc punctiforme		+	+	
Pasteurella multocida	+			
Pseudomonas aeruginosa		+	+	

Psychrobacter arcticus		+	+	
Ralstonia solanacearum		+	+	
Rickettsia conorii				Absent
Rickettsia prowazekii				Absent
Salmonella typhimurium	+			
Sinorhizobium meliloti		+	+	
Slackia heliotrinireductens	+	+	+	Multifunctional fusion protein
Staphylococcus aureus		+	+	
Streptococcus pneumoniae				Absent
Streptococcus pyogenes				Absent
Streptococcus sanguinis		+	+	
Streptomyces coelicolor	+			
Thermoanaerobacter tengcongensis (*Caldanaerobacter subterraneus)*		+	+	
Thermosynechococcus elongatus		+	+	
Thermotoga maritima		+	+	
Treponema caldarium	+			
Treponema pallidum				Absent
Vibrio cholerae	+			
Xanthomonas axonopodis	+			
Xylella fastidiosa	+			
Yersinia pestis	+			

It is not surprising that obligate pathogens like *B. burgdorferi, T. pallidum* and the *Rickettsia* had an absence of *hisG* genes. Host restricted pathogens often have streamlined and highly reduced genomes and are able to scavenge metabolites from within the cell. *E. coli,* and other pathogenic bacteria with free-living stages have significantly larger genomes than host restricted pathogens, and encode 3 – 6 times more proteins (Moran, 2002). With *de novo* histidine biosynthesis being an incredibly energy intensive process, scavenging existing histidine from the intracellular environment of the hosts benefits the pathogen.

Genome reduction in the *Streptococcaceae* has altered histidine synthesis. *S. pyogenes,* and *S. pneumoniae* have an absence of *hisG*, however *S. sanguinis* contains HisG and is likely to generate its own L-His *de novo*. Auxotrophic *Streptococcus* (including *S. pyogenes* and *S. pneumonia)* lack or cannot utilise

genes involved in the pentose phosphate pathway which delivers a precursor ribose-5-phosphate needed for histidine biosynthesis (Willenborg & Goethe, 2016).

The amino acid sequences from species outlined in Table 2.2 were downloaded from InterPro (Blum *et al.*, 2020) and imported into Geneious Prime (version 2019.2.1). Accession codes for each downloaded sequence are included in Appendix 1 (**Table A.1.1**). Using sequence alignment tools in Geneious Prime, Domains I and II were extracted from long or fused forms using reference cores PDB: 5UB9, and 1Z7N. Sequences were then grouped by isoform group; $HisG_L$, $HisG_S$, HisZ, and multifunctional fused forms.

2.2.2 Multiple sequence alignment and maximum likelihood phylogenetic trees

After testing multiple T-coffee alignment programmes (Notredame *et al.*, 2000), T-coffee expresso (version 11.00) (Armougom *et al.*, 2006) was used for the following multiple sequence alignments (MSA). T-coffee expresso MSA was used to create sequence alignments of $HisG_L$ and $HisG_S$ (catalytic core), $HisG_L$, $HisG_S$, and HisZ (four alignments total). Transitive consistency scoring (TCS) was applied using T-coffee workflow after MSA to increase the accuracy of alignments. These MSA were taken to PhyML 3.0 (Guindon *et al.*, 2010) (http://www.atgc-montpellier.fr/phyml/) to build maximum likelihood phylogenetic trees with bootstrapping sampling (100 Bootstrap replicates).

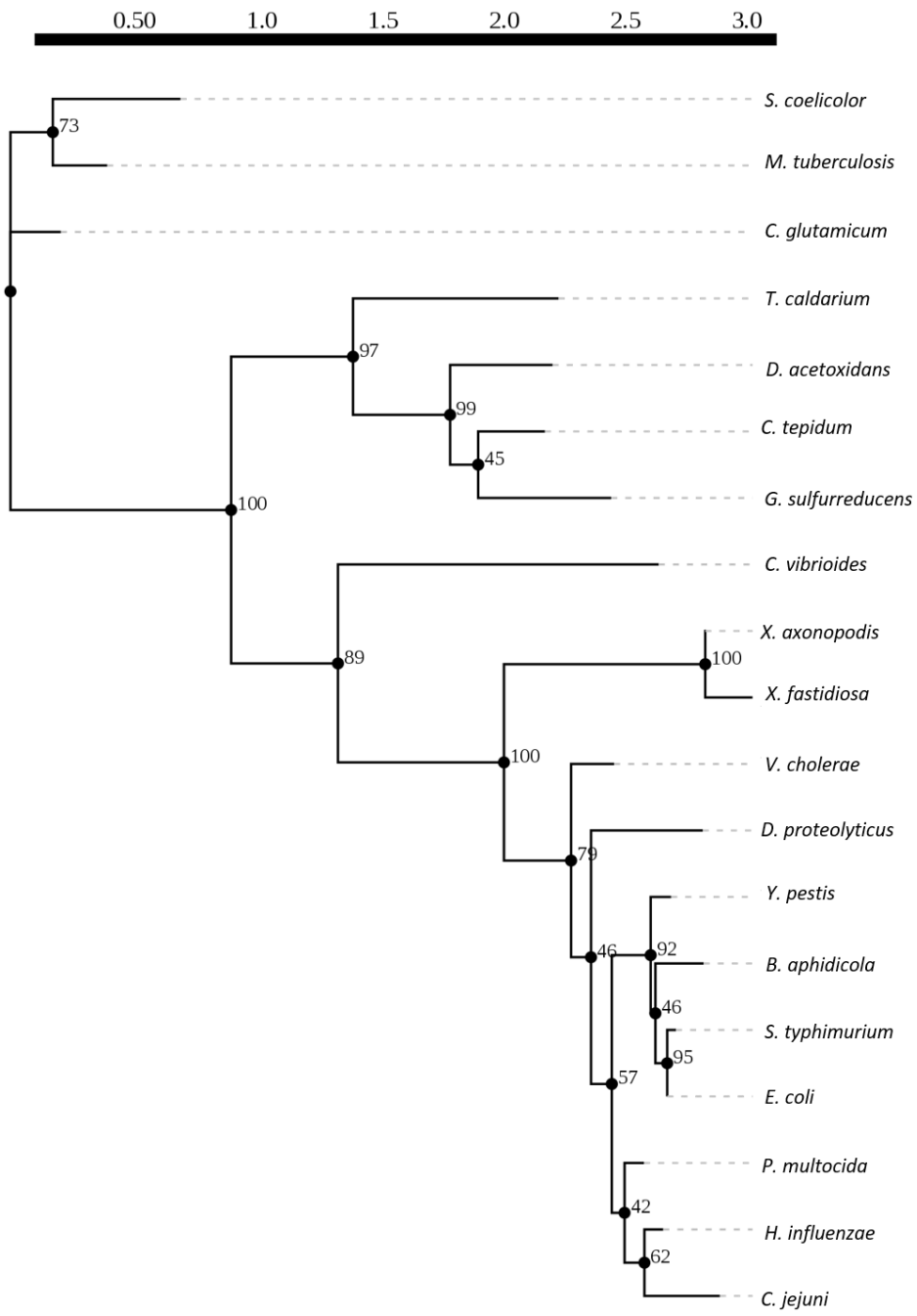

Figure 2.3. Maximum likelihood trees of long form ATP-PRT associated amino acid sequence $HisG_L$. Catalytic core (Domains I and II), and Domain III were aligned using T-coffee expresso, and a maximum likelihood tree generated using PhyML, with 100 bootstrap replicates.

Phylogenetic analysis of the full $HisG_L$ sequence formed three major clades (**Figure 2.3**). *S. coelicolor, M. tuberculosis,* and *C. glutamiucum* (Actinobacteria) formed a distinct clade with deeply rooted branches. The γ-Proteobacteria formed the largest clade with shortest branch lengths. *T. caldarium, D. acetoxidans, C. tepidum,* and

G. sulfurreducens ($HisG_L$) formed a small isolated clade with strong statistical support (bootstrap value of 100).

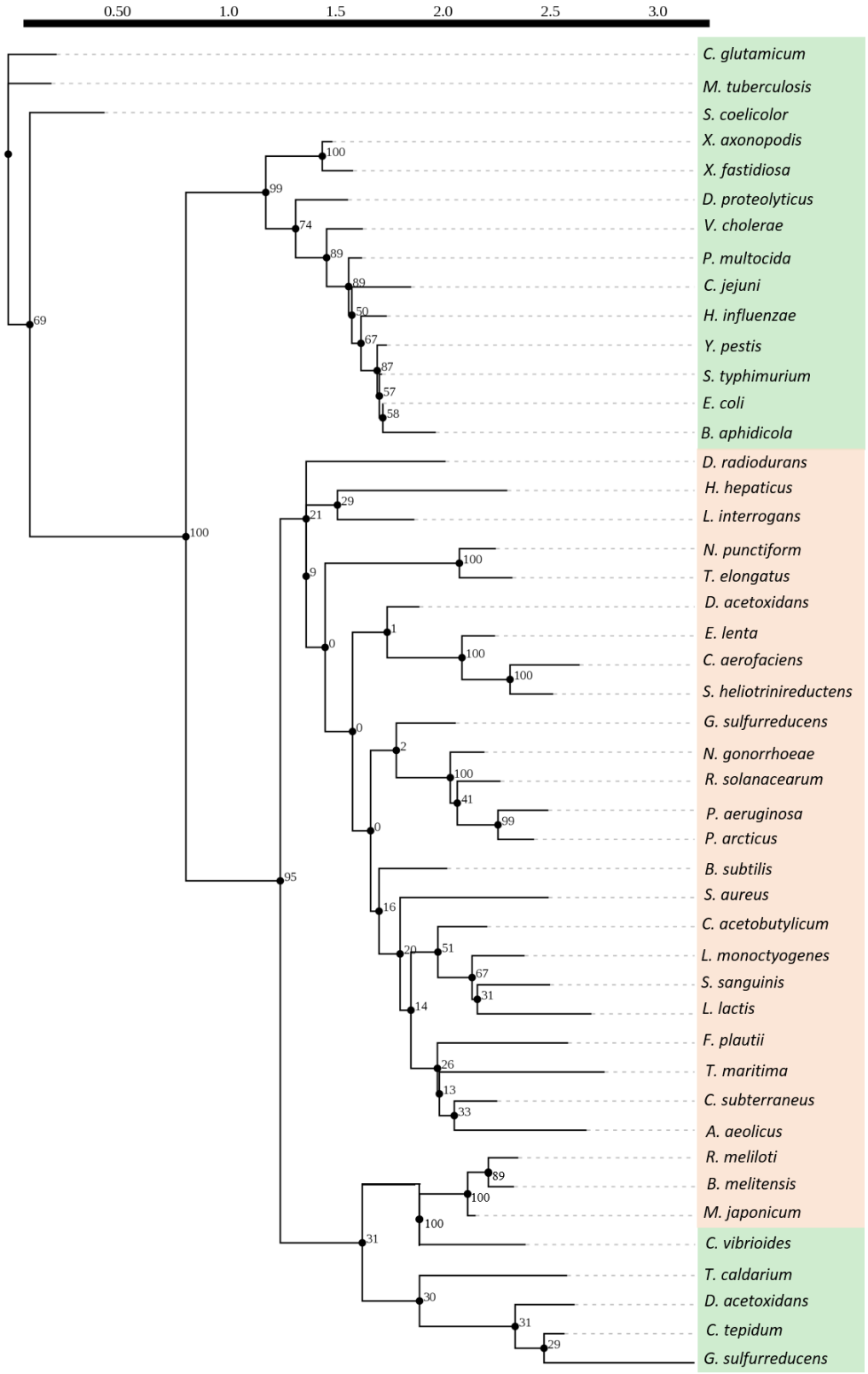

Figure 2.4. Maximum likelihood trees of ATP-PRT associated amino acid sequence. Catalytic core (Domains I and II) of $HisG_L$ were extracted *via* sequence alignment. Multiple

sequence alignment of trimmed $HisG_L$ and $HisG_S$ was performed in T-coffee expresso and a maximum likelihood tree generated using PhyML with 100 bootstrap replicates.

When aligning the shared core domain of $HisG_L$ and $HisG_S$ we see distinct separation between the two genes despite their structural and functional similarity (**Figure 2.4**). Again (like **Figure 2.3**), we see the separation of the Actinobacteria (*S. coelicolor, M. tuberculosis,* and *C. glutamiucum*) forming a distinct clade that is deeply branching. The major clade is broken into two statistically supported groups; the γ-Proteobacteria and a large grouping of organisms containing $HisG_S$. Within the γ-Proteobacteria clade the arrangement of species is very similar to **Figure 2.3** and this clade is well supported by high bootstrap values. In comparison, the large clade of $HisG_S$ has some poor bootstrap values to support further branching in this group. A third clade contains a combination of $HisG_S$ and $HisG_L$. This clade contained $HisG_L$ for the two HisG isozymes, *G. sulfurreducens* and *D. acetoxidans* which grouped with *C. vibrioides, T. caldarium,* and *C. tepidum.* α-Proteobacteria *R. meliloti, M. japonicum,* and *B. melitensis* which were grouped together with high statistical support.

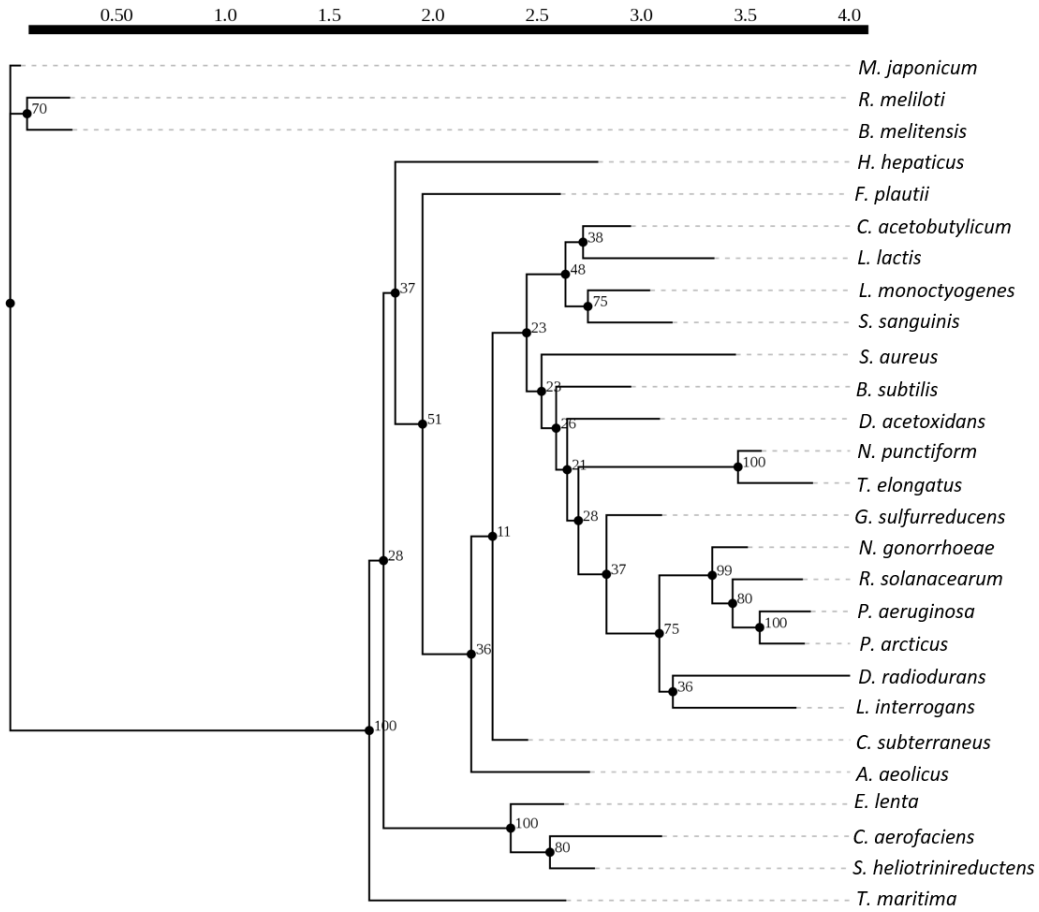

Figure 2.5 Maximum likelihood trees of short-form ATP-PRT $HisG_S$ amino acid sequence. Multiple sequence alignment of $HisG_S$ was performed in T-coffee expresso and a maximum likelihood tree generated using PhyML with 100 bootstrap replicates.

Alignment of $HisG_S$ encoding the catalytic core, domain I and II shows deep branching of α-Proteobacteria *R. meliloti*, *M. japonicum*, and *B. melitensis* which form a distinct and ancient clade (**Figure 2.5**). *T. maritima*, an ancient hyperthermophile is an outlier from the rest of the major clade of $HisG_S$. For the largest clade of $HisG_S$ the branching is not as deep as **Figure 2.3**, which is also seen in **Figure 2.4.**

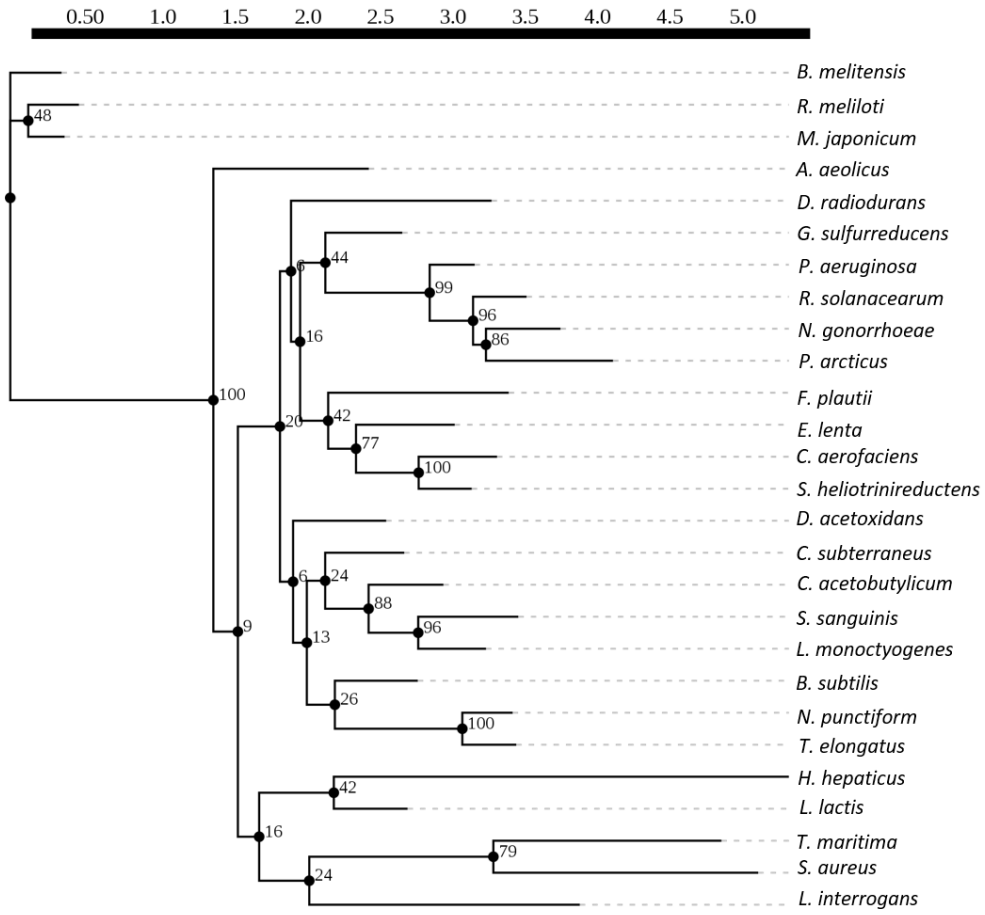

Figure 2.6. Maximum likelihood trees of short-form ATP-PRT regulatory domain HisZ. Multiple sequence alignment of HisZ was performed in T-coffee expresso and a maximum likelihood tree generated using PhyML with 100 bootstrap replicates.

HisZ sequences form a short branched phylogenetic tree (**Figure 2.6**). Like in **Figure 2.4**, the α-Proteobacteria *R. meliloti, M. japonicum,* and *B. melitensis* form a distinct, deep branching ancient clade. Unlike the alignment of *T. maritima* $HisG_S$ (**Figure 2.5**), *T. maritima* HisZ is short branching, and shares a clade with *S. aureus* although the ancient *A. aeolicus* is now an outgroup from this main clade. The remaining species are further arranged in clades similar to their bacterial class.

2.2.3 Gene synteny

To further understand the evolution of APT-PRT isoforms, the operons surrounding HisG were investigated in detail. By exploring the gene synteny in the histidine operon it may be possible to identify regions of possible horizontal gene transfer and generate a hypothesis about regulatory gene transfer. The MicrobesOnline server (Microbesonline.com) (Alm *et al.*, 2005) was used as a search tool for operon

order of the species of interest (**Table 2.2**). These predicted operons were extracted and integrated with phylogenetic trees created in section 2.2.2. Histidine biosynthesis genes A through I, S, and Z were colour coded for observation in operon order (**Figure 2.7**, **B**). Additionally, Adenylosuccinate synthetase (*purA*) was colour coded due to the frequency of which it appeared in operons.

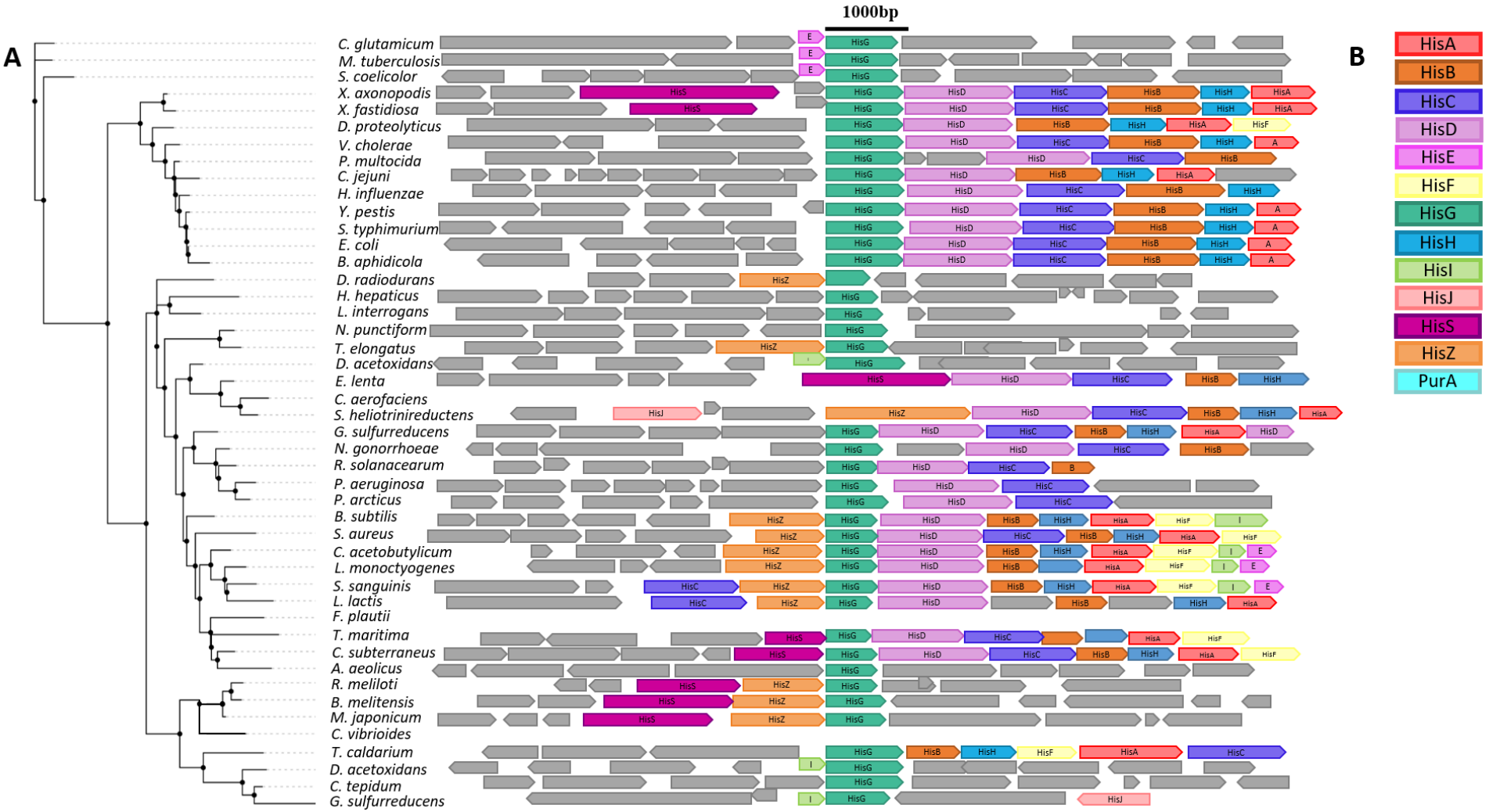
A
1000bp
C. glutamicum
M. tuberculosis
S. coelicolor
X. axonopodis
X. fastidiosa
D. proteolyticus
V. cholerae
P. multocida
C. jejuni
H. influenzae
Y. pestis
S. typhimurium
E. coli
B. aphidicola
D. radiodurans
H. hepaticus
L. interrogans
N. punctiform
T. elongatus
D. acetoxidans
E. lenta
C. aerofaciens
S. heliotrinireductens
G. sulfurreducens
N. gonorrhoeae
R. solanacearum
P. aeruginosa
P. arcticus
B. subtilis
S. aureus
C. acetobutylicum
L. monoctyogenes
S. sanguinis
L. lactis
F. plautii
T. maritima
C. subterraneus
A. aeolicus
R. meliloti
B. melitensis
M. japonicum
C. vibrioides
T. caldarium
D. acetoxidans
C. tepidum
G. sulfurreducens
B
HisA
HisB
HisC
HisD
HisE
HisF
HisG
HisH
HisI
HisJ
HisS
HisZ
PurA

Figure 2.7 A: Phylogenetic tree of bacterial ATP-PRT with associated gene order of genes involved in histidine biosynthesis. Maximum likelihood phylogenetic tree of HisG catalytic core (HisG domain I and II) created by multiple sequence alignment in T-coffee expresso (Armougom *et al.*, 2006), and PhyML tree with PhyML 3.0 (Guindon *et al.*, 2010). The order of histidine biosynthesis associated genes are centred on HisG. Gene order was generated by MicrobesOnline predictive operon server (Alm *et al.*, 2005). Operon size is to scale and coloured by gene (Figure B). *hisG* is not central in the predicted operon for multifunctional fusion proteins *E. lenta* and *S. heliotrinireductens*. MicrobesOnline server identified these as *hisS* and *hisZ* due to gene fusion and labelling has been left as so.

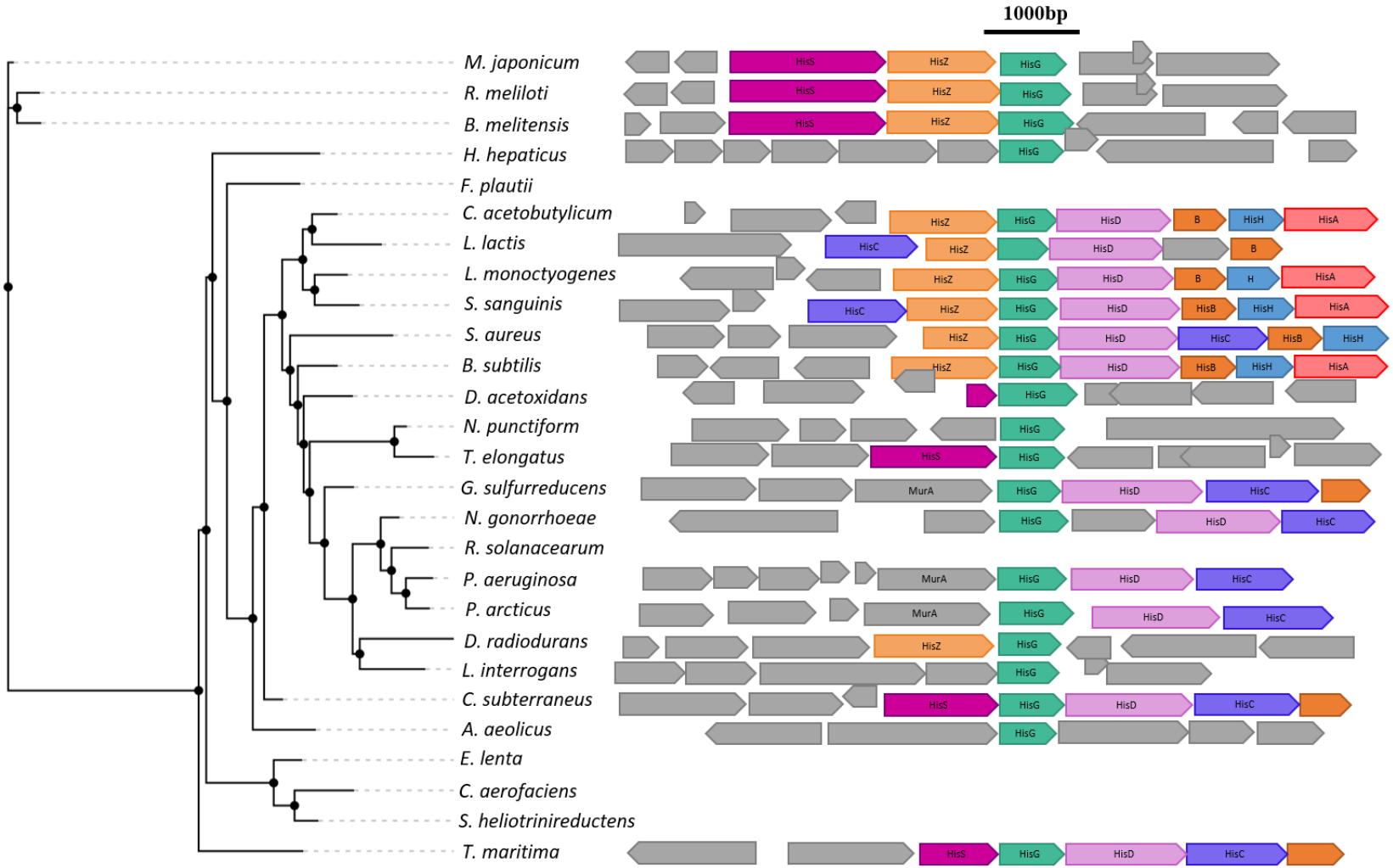
1000bp
M. japonicum
R. meliloti
B. melitensis
H. hepaticus
F. plautii
C. acetobutylicum
L. lactis
L. monoctyogenes
S. sanguinis
S. aureus
B. subtilis
D. acetoxidans
N. punctiform
T. elongatus
G. sulfurreducens
N. gonorrhoeae
R. solanacearum
P. aeruginosa
P. arcticus
D. radiodurans
L. interrogans
C. subterraneus
A. aeolicus
E. lenta
C. aerofaciens
S. heliotrinireductens
T. maritima
HisS
HisZ
HisG
HisD
B
HisH
HisA
HisC
HisB
H
MurA

Figure 2.8. Phylogenetic tree of short-form ATP-PRT with associated gene order of genes involved in histidine biosynthesis. Maximum likelihood phylogenetic tree of $HisG_S$ amino acid sequence created by multiple sequence alignment in T-coffee expresso **(Armougom *et al.*, 2006)** and PhyML tree with PhyML 3.0 **(Guindon *et al.*, 2010)**. Operon synteny centred around $HisG_S$ was generated by MicrobesOnline operon prediction **(Alm *et al.*, 2005)**. Operon size is to scale and coloured by gene (see key **Figure 2.7**; B).


1000bp
B. melitensis
R. meliloti
M. japonicum
A. aeolicus
D. radiodurans
G. sulfurreducens
P. aeruginosa
R. solanacearum
N. gonorrhoeae
P. arcticus
F. plautii
E. lenta
C. aerofaciens
S. heliotrinireductens
D. acetoxidans
C. subterraneus
C. acetobutylicum
S. sanguinis
L. monoctyogenes
B. subtilis
N. punctiform
T. elongatus
H. hepaticus
L. lactis
T. maritima
S. aureus
L. interrogans
HisS
HisZ
HisG
PurA
HisD
HisC
HisB
HisH
HisA
B
H

Figure 2.9 Phylogenetic tree of the regulatory domain of short form ATP-PRT with associated gene order of genes involved in histidine biosynthesis. Maximum likelihood phylogenetic tree of HisZ amino acid sequence from ATP-PRT$_S$ species created by Multiple sequence alignment in T-coffee expresso (Armougom *et al.*, 2006) and PhyML tree with PhyML 3.0 (Guindon *et al.*, 2010). Operon synteny centred around HisZ was generated by MicrobesOnline operon prediction. Operon size is to scale and coloured by gene (see key **Figure 2.7**; B). Where *hisZ* is labled as *hisS* MicrobesOnline server identified these as *hisS* and labelling has been left to preserve the predicted operon information.

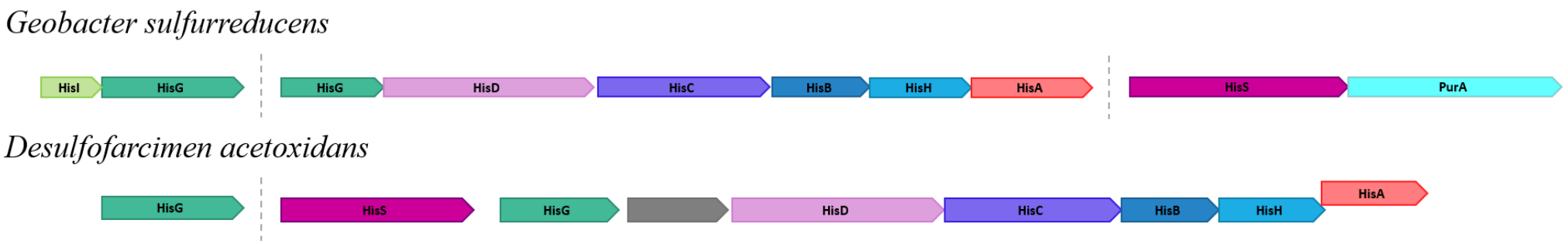

Figure 2.10 Operon order of *hisG* and regulatory genes in ATP-PRT isozymes *Geobacter sulfurreducens* and *Desulfofarcimen acetoxidans*. (see key Figure 2.7**; B**).

Combining the predicted operon order of genes surrounding *hisG* on the phylogenetic tree of catalytic cores from $hisG_L$ and $hisG_S$ we can visualise the way genes move between species, helping to view the evolutionary path (**Figure 2.7**). The three Actinobacter that form a very deeply branched clade, share gene synteny with *hisE* preceding *hisG* but no other evidence of further histidine biosynthetic genes. The γ-Proteobacteria which have shared a clade in each previous phylogenetic tree show fairly high levels of gene synteny, with many containing a number of histidine biosynthetic gene arranged downstream of *hisG*. *P. multicoda* does have an insertion of two genes between *hisG* and *hisD*, which do not function in histidine biosynthesis. There is a selection of short-form species that have the regulatory *hisZ* gene, preceding *hisG* and the subsequent histidine biosynthesis genes.

Examining the predicted operons of $hisG_S$ genes, (**Figure 2.8**) only the α-Proteobacteria *R. meliloti, M. japonicum,* and *B. melitensis*, in a deep-rooted section of the tree, show high operon synteny with *hisS, hisZ, hisG* order, and each respective gene at similar sizes. There is similarly in gene order in the Firmicutes, with HisZ preceding HisG, and followed by subsequent histidine biosynthesis genes. There is also a number of species where $hisG_S$ is either preceded by a short *hisS* (histidyl-tRNA synthetase) or no histidine biosynthesis genes (*H. hepaticus, N. punctiform, A. aeolicus,* and *L. interrogans*). Finally, in four organisms (*G. sulfurreducens*, *P. aeruginosa*, and *P. arcticus*,) $hisG_S$ is preceded by UDP-N-acetylglucosamine 1-carboxyvinyltransferase (MurA) involved in peptidoglycan synthesis.

Figure 2.9 shows operon arrangement surrounding the regulatory gene of ATP-PRT_S *hisZ*. In the majority of organisms where *hisZ* is preceded by $hisG_S$ the operon prediction remains the same. Of interest, is the organisms where $hisG_S$ is fragmented in the genome outside of the histidine operon (*H. hepaticus, N. punctiform, A. aeolicus, G. sulfurreducens, P. aeruginosa, P. arcticus,* and *L. interrogans*), or $hisG_S$ is not preceded or located near *hisZ* (*D. acetoxidans, T. elongatus, N. gonorrhoeae, C. subterraneus,* and *T. maritima*). Of these organisms, *A. aeolicus, G. sulfurreducens, C. subterraneus, H. hepaticus, T. maritima,* and *L. interrogans hisZ* is not differentiated from *hisS* and as a result for *C. subterraneus* and *T. maritima hisZ* may directly precede $hisG_S$. The most

notable trend in the predicted genes surrounding *hisZ* is the association with *purA*. In a number of species where regulatory *hisZ* is not found near $hisG_S$, it is instead located in another genomic region neighbouring *purA*, which encodes an enzyme in the purine biosynthesis pathway.

Figure 2.10 illustrates the predicted operon order of ATP-PRT isozymes *G. sulfurreducens* and *D. acetoxidans*. Containing both $hisG_S$ and $hisG_L$ we see for both organisms $hisG_L$ is not associated with an operon of histidine biosynthesis genes. For *G. sulfurreducens* $hisG_L$ is preceded *hisI* (step two and three of histidine biosynthesis). *G. sulfurreducens* $hisG_S$ is followed by HisDCBHA, but regulatory gene *hisZ*/*S* is located elsewhere in association with *purA*. For *D. acetoxidans* $hisG_L$ is located outside of any *his* operon, and $hisG_S$ is located in an operon between *hisS* and HisDCBHA.

2.2.4 List of bacteria with *hisZ*/*purA* gene association

Figure 2.9 illustrates an unexpected pairing of *hisZ* and *purA* observed across operons of $hisG_S$. Following the building of these trees, MicrobesOnline was used to determine the scope of *hisZ*/*purA* gene association across bacterial genuses, that were otherwise excluded from previous phylogenetic trees.

Table 2.3. List of Genera containing an operon order in which *purA* and *hisZ* are neighbouring as determined using MicrobesOnline operon and regulon server (Alm et al., 2005).

Acidiphilium	*Cupriavidus*	*Methylibium*	*Rhodoferax*
Acidithiobacillus	*Dechloromonas*	*Methylobacillus*	*Rhodospirillum*
Acidovorax	*Diaphorobacter*	*Methylophilales*	*Saccharophagus*
Acinetobacter	*Geobacter*	*Methylotenera*	*Sideroxydans*
Alcanivorax	*Gluconobacter*	*Methylovorus*	*Sorangium*
Alkalilimnicola	*Hahella*	*Neisseria*	*Sphingobium*
Allochromatium	*Halorhodospira*	*Nitratifractor*	*Teredinibacter*
Azoarcus	*Halothiobacillus*	*Nitrococcus*	*Thauera*
Azospirillum	*Helicobacter*	*Nitrosococcus*	*Thioalkalivibrio*
Azotobacter	*Herminiimonas*	*Nitrosomonas*	*Thiobacillus*
Bordetella	*Hippea*	*Nitrosospira*	*Thiomicrospira*
Burkhoderia	*Janthinobacterium*	*Oceanobacter*	*Thiotrix*
Caminibacter	*Laribacter*	*Oceanospirillum*	*Variovorax*
Campylobacter	*Leptothrix*	*Plesiocystis*	*Verminephrobacter*
Candidatus	*Limnobacter*	*Polaromonas*	*Wolinella*
Chromobacterium	*Magnetospirillum*	*Polynucleobacter*	*Xanthobacter*

Chromohalobacter	*Marinobacter*	*Pseudomonas*	*Zymomonas*
Comamonas	*Marinomonas*	*Psychrobacter*	
Congregibacter	*Mariprofundus*	*Ralstonia*	
Delftia	*Methylococcus*	*Reinekea*	

2.2.5 PFAM search

Histidine biosynthesis is widely spread across bacteria, however the low sequence similarity of *hisG* genes makes creating a universal tree difficult. Instead PFAM (Mistry *et al.*, 2020) was used to explore the spread of *hisG* genes across microbes. Using PFAM accessions for *hisG* (PF01634) and HisG_C (PF08029) starburst maps were generated for *hisG* containing species and HisG_C containing species (**Figure 2.11**). Unfortunately, due to the sequence similarity between *hisZ* and tRNA-synthetases there is no specific PFAM accession code to retrieve *hisZ* from all microbes. The generated starburst maps for HisG and HisG_C illustrate the distribution of each gene across different microbial groups. This provides a visual tool to investigate the spread of *hisG* genes in different bacterial groups, and the prevalence of each respective gene.

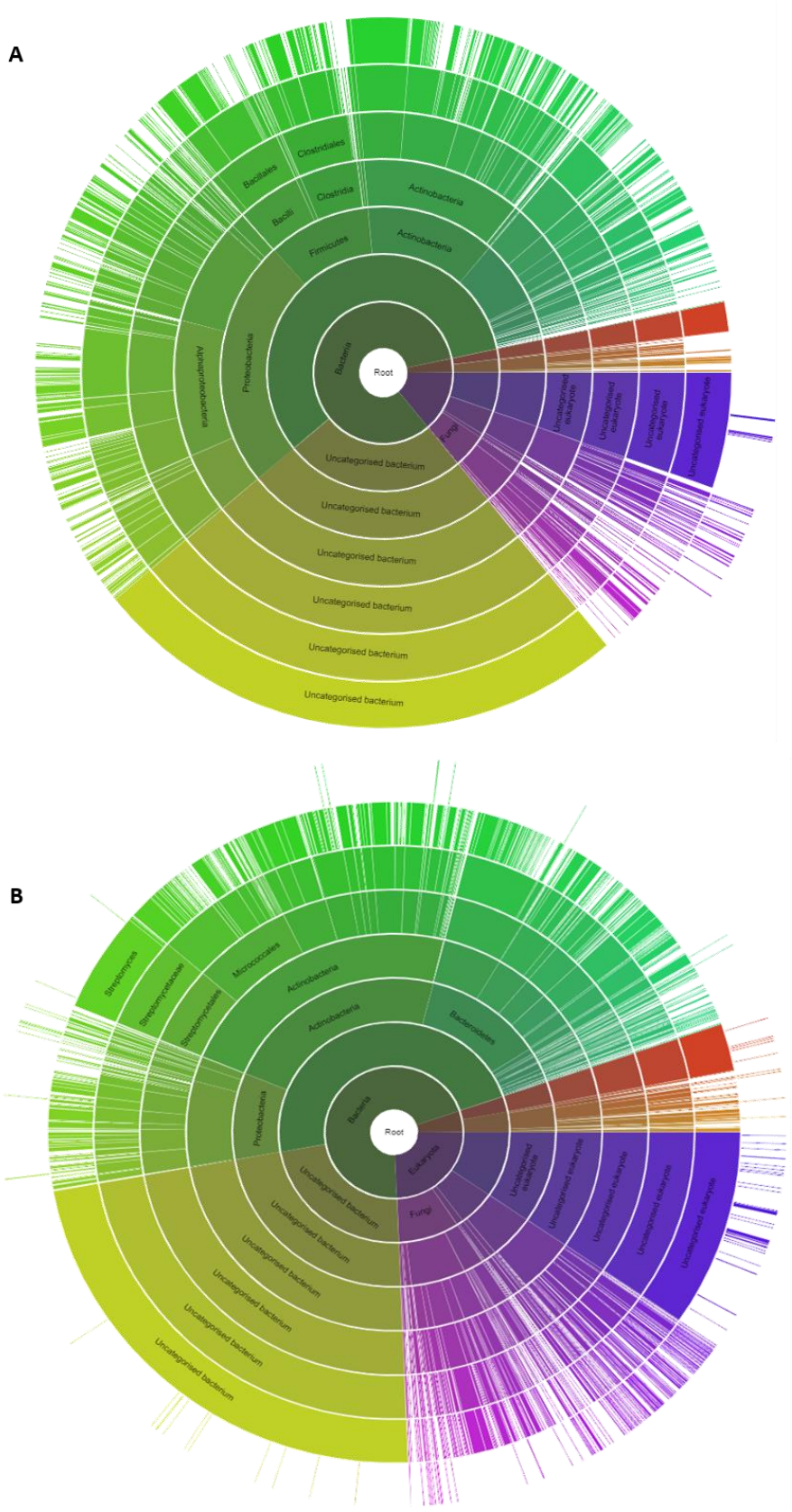

Figure 2.11. Starburst map of *hisG* associated genes across species. A: Starburst map of species containing *hisG* (PFAM accession: PF01634). B: Starburst map of species containing HisG_C (PFAM accession: PF08029). Branches coloured by domain; Bacteria; green, Eukarya; purple, and Archaea; displayed in orange.

Based on **Figure 2.11**, 6648 bacterial species contained PF01634 (*hisG*), and 3068 bacterial species contained HisG_C (PF08029). Of particular note, the Firmicutes bacterial group did not account for any of the bacterial species containing HisG_C, despite making up 12.4% of all the species.

2.3 Discussion

The two ATP-PRT isoforms have long been a topic of evolutionary investigation, as two sophisticated regulatory methods from one enzyme of core metabolism is intriguing. In this chapter, traditional phylogenetic trees using protein sequences of bacterial *hisG* along with operon analysis have been used to investigate trends in *hisG* distribution, the preservation of gene order and the possible evolutionary trajectory for these genes.

2.3.1 Multiple sequence alignments and maximum likelihood trees

T-coffee expresso created MSA of the catalytic core of selected bacterial species, and alignments for the amino acid sequences of both $HisG_S$ and HisZ of short-form ATP-PRT. The protein sequence identity of all ATP-PRT's is very low, from 3.3, 5.5, 6.2 % identical sites of the core domain shared between isoforms, in $HisG_L$ (full sequence) $HisG_L$ (core), and $HisG_S$ respectively, suggesting a long divergence period since the last common ancestor of either group. Prior to MSA using T-coffee, sequence alignments were performed in Geneious Prime. Pairwise sequence identity between even the same isoforms was below 25 %, a factor that made the phylogenetic analysis challenging. Maximum likelihood trees created in PhyML from multiple sequence alignments generated in T-coffee expresso created three phylogenetic trees that incrementally investigated the regulatory domains of bacterial ATP-PRT.

2.3.1.1 Distribution of $HisG_L$ and $HisG_S$

The initial trees (**Figure 2.3**, **Figure 2.4**, and **Figure 2.5**) which aligned core domains created few major branches. The largest branch contained groupings of $HisG_S$, with the second largest containing $HisG_L$ core. The distribution of long and short HisG isoforms from the MSA and PhyML tree using the core sequence

indicates a segregation of the isoforms even with the catalytic core being well conserved both functionally and structurally. Excluding *T. caladarium*, *D. acetoxidans* (long), *G. sulfurreducens* (long), and *C. tepidum*, $HisG_L$ formed a monophyletic group. This suggests $HisG_L$ and $HisG_S$ diverged into strict groups early in prokaryotic evolution. *B. melitensis*, *S. meliloti*, and *M. japonicum* formed deeply branched outliers on each phylogenetic tree (**Figure 2.5** and **Figure 2.6**), these three species also formed a small group in the phylogenetic tree by Battistuzzi *et al.* (2004) that branched off from the long-form group (**Figure 2.2**). *B. melitensis* and *S. meliloti* are noted to share orthologues of a number of genes (Jones *et al.*, 2007; Mirabella *et al.*, 2013) and have possibly co-evolved histidine biosynthetic genes over time.

Short-form ATP-PRT (**Figure 2.5**, and **Figure 2.6**) shows shuffling of clades between trees, indicating independent gene transfer events of the core and HisZ domains. While the majority of species retained similar clades between the $HisG_S$ and HisZ trees, *D. acetoxidans*, *B. subtilis, L. lactis,* and *T. maritima* appear scattered. This dissimilarity between the two short-form trees shown in **Figure 2.5** suggests in *D. acetoxidans, B. subtilis, L. lactis,* and *T. maritima* HisZ and $HisG_S$ were not inherited together. If both gene products had been simultaneously introduced, thus sharing the same common ancestor, the homogeneity of gene location and clades should remain similar. $HisG_S$ of *B. subtilis* is within a clade containing *S. sanguinis, L. monocytogenes, C. acetobutylicum, and C. subterranus,* however the HisZ of *B. subtilis* outgroups with isozyme *D. acetoxidans*. This indicates that gene swapping of the two domains independently has been a common occurrence after the establishment of the HisGs – HisZ cooperative unit.

Gene rearrangement across these histidine operons can be attributed to horizontal gene transfer (HGT), recognised as playing a significant role in the development of histidine biosynthesis and its regulation (Bond & Francklyn, 2000; Fani *et al.*, 2005). Histidine genes and operon segments have likely been horizontally transferred between bacteria and proteobacterial branches creating scattering throughout the genome, resulting rearrangement to the phylogenetic trees produced. Operons, where present, allow for compaction of genes, co-regulation of protein synthesis, and the generation of enzymes in ideal ratios. However the scattering of histidine genes in these, and other organisms, supports evidence of multiple HGT events or

operon assembly occurring multiple times over long periods (Fondi *et al.,* 2009). The histidine synthesis and regulatory genes in *N. gonorrhoea* are scattered throughout the genome. This scattering is indicative of HGT, which can be expected in naturally competent organisms. *P. arcticus, P. aeruginosa, R. solanacearum, and H. hepaticus* are also recognised to have natural competency, and we see they also show fragmented histidine operons (**Figure 2.10**).

2.3.1.2 Distribution of multifunctional fused HisG and HisZ

Within the major branch of $HisG_S$ (**Figure 2.5**) there is a monophyletic group dedicated to the multifunctional fused $HisG_S$/HisZ sequences of *E. lenta*, *C. aerofacien*s, and *S. heliotrinireductens.* These three species remain grouped together across all sequential trees (**Figure 2.5**) due to their high sequence identity (36.4% identical sites). Each species belongs to the *Coriobacteriaceae* of the Actinobacteria (Clavel *et al.,* 2014). UniProt shows 106 entries of *Coriobacteriaeceae hisG* genes, many of these are 500 residues in length indicating additional multifunctional fused *hisG/hisZ* isoforms. This fusion being limited to the *Coriobacteriaceae* genus signifies a recent fusion of genes, and indicates the multifunctional form is not ancestral. However, the fusion of *hisG*/*hisZ* allows tight co-regulation of the two gene products, essential for ATP-PRT_S function.

2.3.1.3 Distribution of HisG isozymes

D. acetoxidans and *G. sulfurreducens* are two bacterial species that are recognised to contain both functional $HisG_L$ and $HisG_S$. Based on the initial phylogenetic trees (**Figure 2.4**) $HisG_L$ and $HisG_S$ from both species are scattered across the tree and branch off or near groupings of other $HisG_L$ and $HisG_S$ cores respectively. $HisG_S$ from both *D. acetoxidans* and *G. sulfurreducens* branches off within the main branch of $HisG_S$ cores, indicating similarity to modern $HisG_S$ genes. $HisG_L$ core, however, diverges early on, forming a deeply branching clade with *T. caldarium* and *C. tepidum*. In both *D. acetoxidans* and *G. sulfurreducens*, the analogous out group of $HisG_L$ suggests its ancient nature, suggesting it may be the older form.

When investigating these isozymes using accompanying operon order, it is evident that in *G. sulfurreducens* $hisG_S$ sits within an ordered histidine operon (HisZGDBHAFI). In *D. acetoxidans* a similar operon structure is seen

(HisSGDCBHAFI). In both *G. sulfurreducens* and *D. acetoxidans* $hisG_L$ is accompanied only by a truncated *hisI*. This supports earlier evidence that $hisG_L$ is the older isoform in these species, and may be less essential to the survival of the bacteria, evident by knockdown studies by (Aklujkar, 2011). Isozymes are usually present as a result of gene duplication, when these multiplied genes persist over time. In the case of *G. sulfurreducens* and *D. acetoxidans* when both isoforms persist over time, we can assume they both serve to regulate in unique and beneficial ways, perhaps optimised under different environmental conditions.

2.3.1.4 How old is *hisG*?

This research did not identify an age to either of the *hisG* genes, nor an ancestral form. The sequences for each histidine biosynthesis gene had low sequence similarity, and preliminary trees generated, produced poor statistics and thus would be unsuitable for dating. Fondi *et al.* (2009) identified histidine biosynthesis as an ancient pathway that likely links back to before the divergence of Bacteria, Archaea, and Eukarya. In their attempts to identify the last universal common ancestor (LUCA) they were also faced with low bootstrap statistics. The present analysis confirms histidine biosynthesis to be an ancient gene pathway, evident by very low sequence similarity of *hisG*.

2.3.2 HisG_C is absent from the Firmicutes

A surprising discovery from this research was the complete absence of HisG_C from the Firmicutes, identified while investigating PFAM starburst maps (**Figure 2.11**). All Firmicutes included in the MSA and resulting phylogenetic trees were short-form and therefore lacking HisG_C. This observation illustrates the distribution of HisG_C and by association $HisG_L$ across a large proportion of the bacterial tree. However, upon further investigation, UniProt returned two Firmicute species containing probable $hisG_L$; *Desulforudis audaxviator* and *Pelotomaculum thermopropionicum*, with *hisG* lengths of 289 residues. By aligning the protein sequences of both ATP-PRT species, these species shared 46.9 % identical sites with $hisG_L$ from *G. sulfurreducens*. For both species, $hisG_L$ is seemingly randomly placed in the genome, outside of any histidine operon (**Figure 2.12**). The high sequence similarity between these genes, and placement outside of a histidine operon is suggestive of a recent HGT event. HisDCBHAFI are located

elsewhere in the genome, although conserved compared to other predicted operons in this study these genes are located next to PurEBCFMNHD of purine biosynthesis (note PurEBCF are not included in **Figure 2.12** due to length).

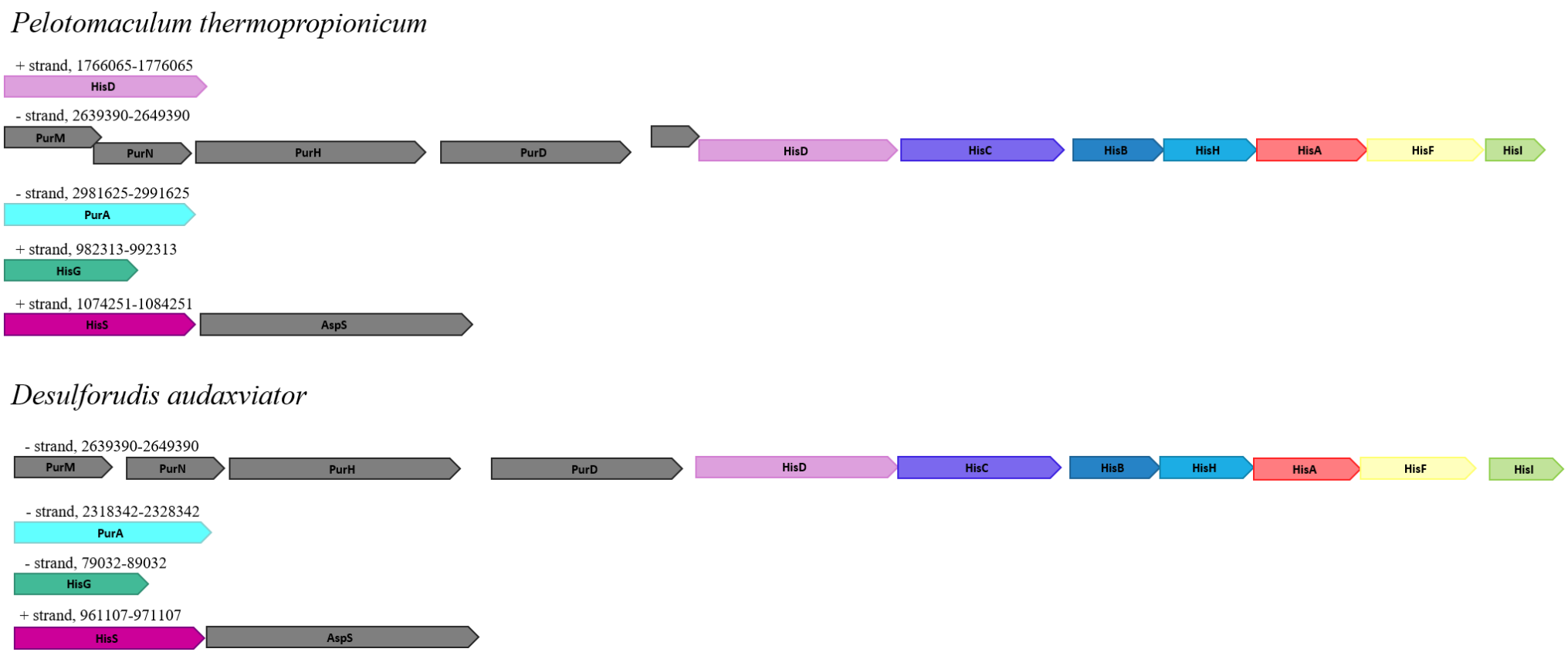

Figure 2.12 Predicted operons of histidine biosynthesis associated genes in *Pelotomaculum thermopropionicum* and *Desulforudis audaxviator* (Predicted operon order from microbesonline.com). Both organisms show similar gene synteny of Histidne operon, however, *hisG*$_L$ is located outside of the histidine operon in both species suggesting it has been introduced in a separate event compared to hisA-DFHI.

2.3.3 *hisG$_L$* is likely the ancestral form

Low sequence similarity between HisG$_L$ and HisG$_S$, identified through multiple sequence alignment, indicates the divergence of the enzyme sequence, and thus indicates the old age of the enzyme and likely the whole histidine biosynthetic pathway. As a result of the low sequence similarity driven by long periods of gene divergence, phylogenetic approaches using sequence alignment and phylogenetic trees alone were unable to definitively determine the ancestral form of HisG. Nonetheless, by combining phylogenetic trees with predicted operons (**Figure 2.7**, **Figure 2.8**, **Figure 2.9**, and **Figure 2.10**), and global presence and absences of *hisG* gene variants (**Figure 2.11**) we were are able to postulate *hisG$_L$* as the ancestral form. Due to the vast spread of HisG_C across all bacterial phylogenies bar the Firmicutes, we suspect *hisG$_L$* as the ancestral form of all regulated ATP-PRT isoforms.

The ancestral *hisG$_L$* dispersed among the three domains of life (**Figure 2.11**), evident by the distribution of the C-terminus across all three domains (Bacteria, Archaea, and Eukarya). *hisG$_L$* evolved and diverged within the bacteria with the complete loss of C-terminus (in the Firmicutes). The Firmicutes are the only distinct phylum group missing HisG_C (except *D. audaxviator* and *P. thermopropionicum*).

The dispersal of *hisG$_L$* across each domain, along with the deep branching clades of HisG$_L$ adds evidence of HisG$_L$ being the ancestral form. For *hisG$_S$*, the regulatory gene *hisZ* bares clear similarity to existing histidyl-tRNA synthetases (*hisS*), additionally, aminoacyl-tRNA synthetases (aaRSs) are prone to HGT (Wolf *et al.*, 1999). It is then reasonable to predict an ancient aaRSs ancestor of *hisZ* became inserted and co-opted as a regulatory domain, followed by the subsequent loss of the C-terminal regulatory domain (HisG_C).

The phylum Firmicutes, is the only distinct group that predominantly uses the HisZ regulatory domain. The gene synteny of histidine operon within the Firmicutes is well conserved with *hisZ* preceding the short-form *hisG* (**Figure 2.8**). This operon conservation, and distinct lack of HisG_C suggests the LUCA of Firmicutes contained the co-opted aaRSs, and had already undergone the loss of HisG_C.

The newly co-opted regulatory domain would then spread through the bacterial tree *via* HGT, resulting in the disordered gene synteny of *hisG*$_S$, and infrequent coordination of *hisG*$_S$ and regulatory *hisZ*. The gene synteny of other bacterial short-form ATP-PRTs is not as well conserved and instead, many regulatory *hisZ* genes were prominently associated with *purA* (**Figure 2.9**), suggesting multiple HGT events occurred resulting in disordered operons. Further comparison of operon diversity, and calculated divergence of *hisZ* from *hisS* or aaRSs could offer information into the acquisition of *hisZ* as a regulatory domain in *hisG*$_S$.

2.3.4 Histidine and Purine biosynthesis linkage

An unexpected finding of this evolutionary analysis was the linkage of *hisZ* and *purA* within the same predicted operon. Predominant *hisZ/purA* synteny in bacteria suggests co-transcription of the two genes or may link the two gene products for biosynthesis. The roles of *hisZ* and *purA* can be interlinked purely by the crossover of histidine and purine biosynthesis via 5-aminoimidazole-4-carboxamide ribonucleotide (AICAR) **Figure 2.13**. Note that *purA* encodes is the first committed step of ATP synthesis and is downstream of AICAR in the synthesis of ATP which itself, is a precursor for histidine. Thus, it makes sense for *purA* and *hisZ* to be co-regulated at a transcriptional level.

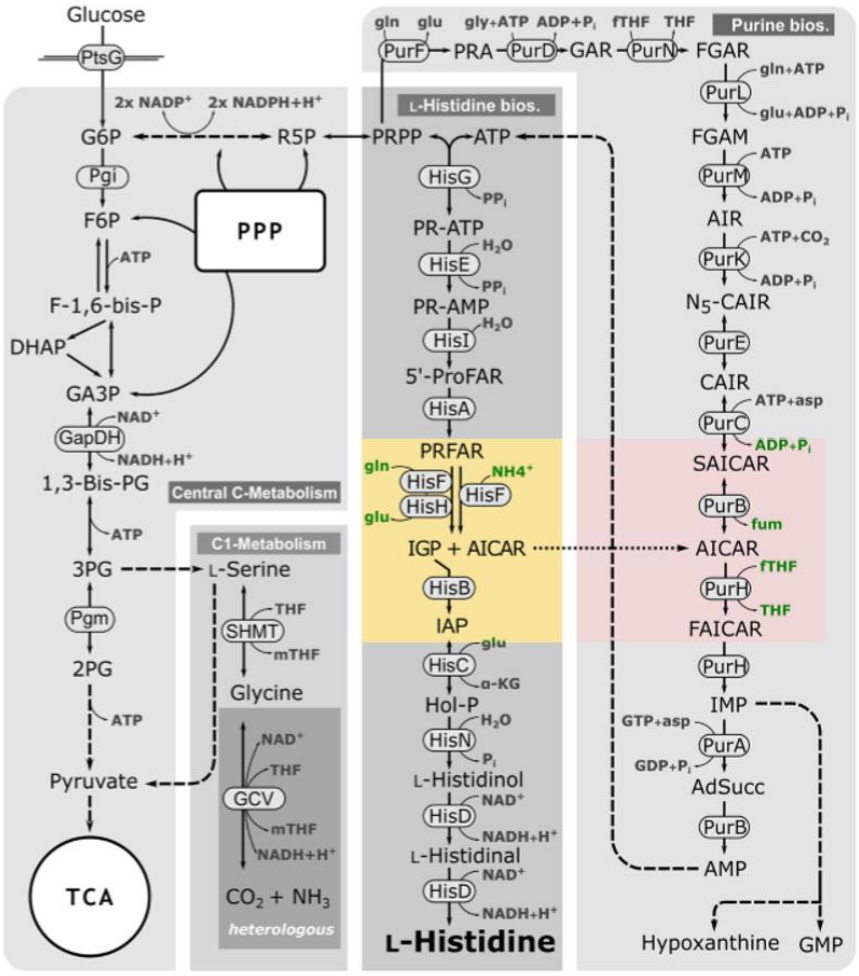

Figure 2.13 Interlinking metabolic pathways of L-His synthesis, purine biosynthesis, and central C-metabolism adapted from (Schwentner et al., 2019). AICAR (coloured) leaves histidine biosynthesis as a bi-product at the fifth step. AICAR can then be recycled into ATP *via* purine biosynthesis salvage pathway, therefore replenishing initial substrate.

Histidine biosynthesis is a complex metabolic network, making the overproduction of histidine difficult. Histidine and purine biosynthesis are interlinked, and both require PRPP as a substrate. Additionally, histidine biosynthesis requires ATP in high levels for efficient histidine production. AICAR is formed in the fifth step of histidine biosynthesis and is then rerouted into a four-step purine biosynthesis pathway which can then regenerate the ATP required for histidine biosynthesis.

Although ATP-PRT is denoted as a rate limiting enzyme for histidine biosynthesis in *Corynebacterium glutamicum* (*cgl*ATP-PRT), even when the regulatory domain (HisG_C) of *cgl*ATP-PRT is removed histidine fails to be over produced. In *C. glutamicum*, Schwentner *et al.* (2019) observed that histidine biosynthesis encountered a bottleneck after the fifth step, with AICAR accumulation, and low

ATP. Additionally, inosine monophosphate (IMP) build-up within the AICAR salvage pathway appears to inhibit histidine production in *C. glutamicum*. PurA alleviates IMP build-up and catalyses the third step of ATP salvage (Schwentner *et al.*, 2019). Through upregulation of PurA and simultaneously truncating the ATP-PRT regulatory domain Schwentner *et al.* (2019) were better able to over express L-His in *C. glutamicum*. This interconnection supports a strong feedback link between the two pathways of L-His and purine synthesis.

2.3.4.1 Is AICAR a competitive inhibitor of ATP-PRT?

E. coli ATP-PRT is sensitive to inhibition by AICAR which acts as an analogue of AMP. AICAR is generated equimolar to histidine, therefore the liberation of AICAR would be necessary for continuous production of histidine. This liberation step would also function toward substrate renewal, refreshing the pool of available ATP. The overexpression of PurA and PurH (bifunctional AICAR transformylase/IMP cyclohydrolase) was found to be imperative for readjustment of AICAR and IMP levels (Malykh *et al.*, 2018).

Potentially, the role of AICAR as a competitive inhibitor in these species could drive the co-regulation of purine salvage genes, in a response to liberate competitor molecules inhibiting histidine synthesis. If AICAR is created in equimolar concentrations to histidine, it may function as an effector molecule, decreasing PR-ATP production by ATP-PRT, thus regulating entry point histidine production (**Figure 2.14**).

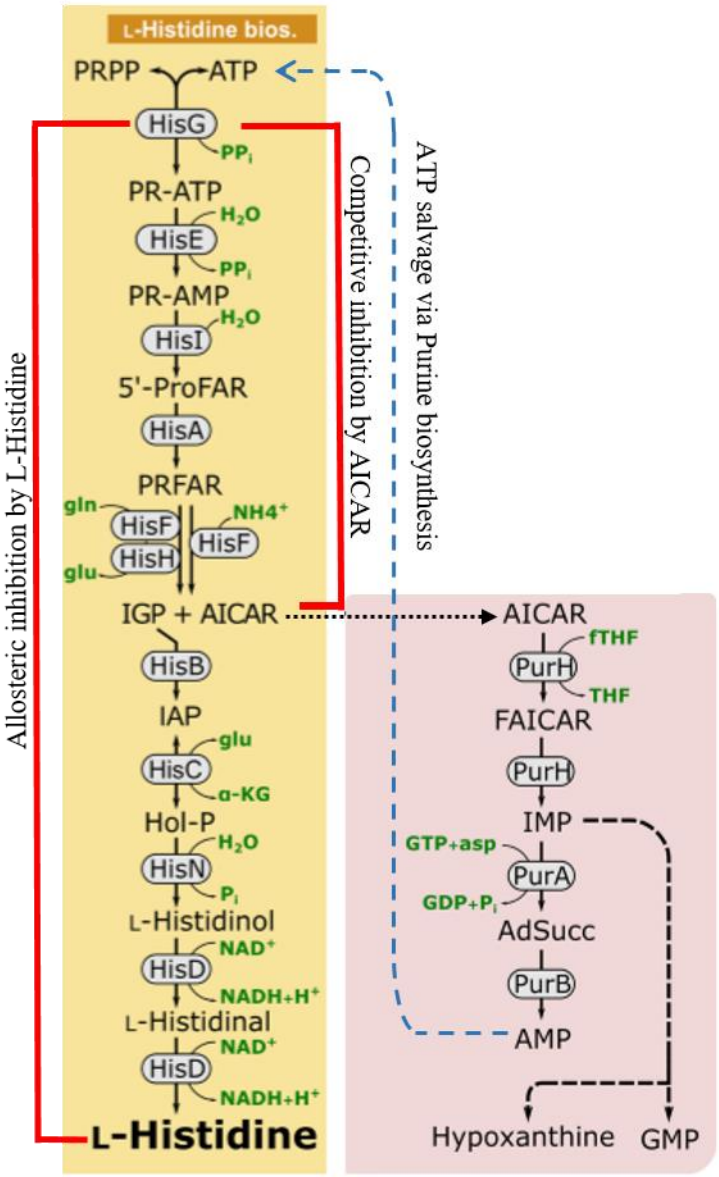

Figure 2.14 Linked histidine and purine biosynthetic pathways adapted from Schwentner et al. (2019). Highlighting pathways of allosteric inhibition by histidine, and possible pathways of competitive inhibition by AICAR and possible substrate salvage pathways.

Reducing AICAR and IMP, and recycling to ATP, can overcome the bottleneck of AICAR thereby increasing histidine production in some bacterium. PurH and PurA are necessary to increase AICAR conversion into ATP. In *C. glutamicum* with both PurA and ATP-PRT controlling rate regulating steps in each metabolic pathway, the conversion of IMP by PurA relieves the build-up of IMP and allows the production of ATP which may feedback to the first step of histidine synthesis thus increasing histidine production (Kulis-Horn *et al.*, 2014). In *E. coli* AICAR can

competitively inhibit ATP-PRT, thus the immediate liberation of AICAR by PurH, allows the continuation of histidine biosynthesis whilst accelerating purine replenishment (Malykh *et al.*, 2018).

2.3.4.2 Histidine biosynthesis is part of central metabolism

In sequential gene inactivation in *Acinetobacter baylyi* (*A. baylyi*), histidine synthesis was noted as the seventh most essential gene set necessary for survival (Gallagher *et al.*, 2020). Sequential mutations were made in many gene sets and time was recorded to assess when growth was inhibited, and thus how essential the gene set was. Histidine synthesis ranked seventh in the list of essential genes (**Figure 2.15**), above ATP synthase, dNTP synthesis, cytochrome oxidase, ribosomal proteins, NADH dehydrogenase, pentose phosphate shunt. Histidine synthesis was additionally ranked higher than other amino acid biosynthesis genes (Gallagher *et al.*, 2020).

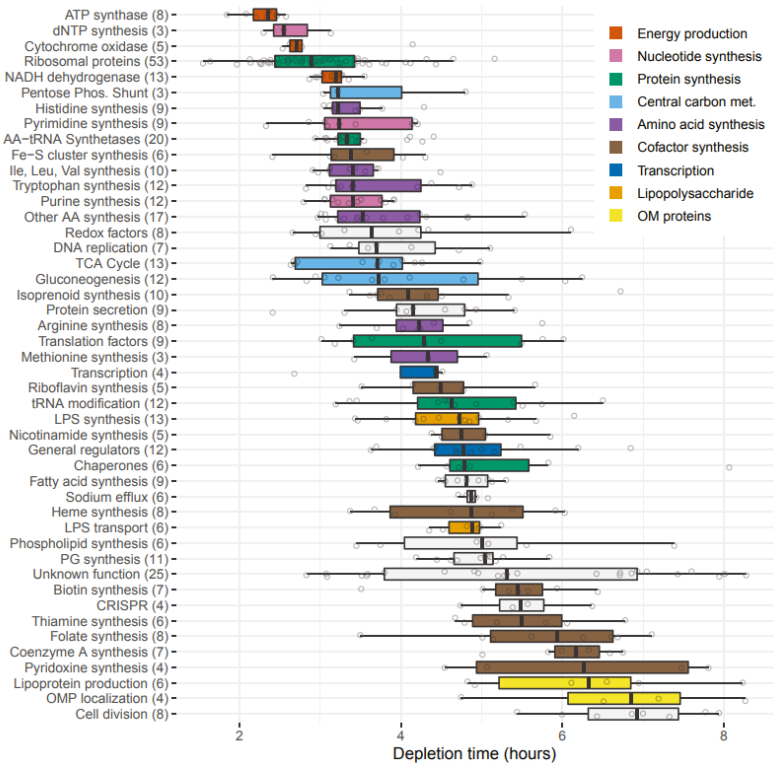

Figure 2.15 Mutation depletion times for essential biological processess (Gallagher et al., 2020). Note histidine synthesis ranked higher than any other amino acid and purine synthesis gene set. Figure used with permission from PNAS.

The position of histidine synthesis genes compared to other *de novo* amino acid biosynthesis genes, could hint at histidine biosynthesis being involved in energy synthesis or possibly further into purine synthesis. Energy and purine synthesis were both highly essential metabolic functions in the sequential gene inactivation. Therefore, this pathway may not function for the sole purpose of histidine production, but as a product in an intertwined metabolic network vital for cell survival.

2.3.4.3 Origin of *hisZ/purA* gene association

With a significant proportion of *hisZ/purA* association, and evidence of purine and histidine linkages apparent in *P. thermopropionicum* and *D. audaxviator,* this gene association could be an artefact of HGT. For numerous species, like the *N. gonorrheae, P. acrticus,* and *P. aeruginosa,* HisGDC appear harmonious, whereas *hisZ* is seemingly randomly placed alongside *purA.* There is evidence of some multifunctional fused *hisZ/purA* genes in the Proteobacteria and Burkholderiacea (UniProt Entry; A0A520VWG8). Could this suggest *hisZ* and *purA* were originally placed together and have distributed among bacteria *via* HGT?

An adenylosuccinate synthetase would benefit from assistance of an ancestral tRNA-synthetase, one gene could assist in the gene regulation of the other. Other Pur genes (*purR*) have been known to have secondary functions as regulators in other biological functions, such as virulence in *Staphylococcus aureus* (Sause *et al.*, 2019). There is a possibility that an ancestor to *hisZ* could function as a transcriptional regulator of *purA* and thus they were located together in an operon, and this gene association was retained upon HGT into other species.

2.3.4.4 Obligate bacterial pathogens retain histidine biosynthesis

Due to the nature of host restricted pathogens these organisms have streamlined their genomes and scavenge nearly all of the metabolites they need from the host. As a result of this metabolite scavenging, many of the host restricted pathogens present the phylogenetic tree by Battistuzzi *et al.* (2004) are histidine auxotrophs, and therefore do not contain *hisG* and associated regulatory genes.

N. gonorrhoeae, an obligate human pathogen is an interesting case of a species that has retained the histidine biosynthesis genes, despite likely being able to scavenge histidine requirements from the host. Transcriptional landscape and knockout studies showed that of all the histidine genes in *N. gonoorrhoeae,* only *hisE* (Step two and three of histidine biosynthesis) was identified as essential for bacterial growth (Remmele *et al.*, 2014). The defined media used by Remmele *et al.* (2014) contained L-His, which may explain why *hisG* is not included in the essential genes list.

N. gonorrhoeae, is capable of up taking a number of essential amino acids from the environment, including L-His (Catlin, 1973). If uptake from the intracellular environment is possible, what factors are driving *N. gonororrhoeae* to retain *de novo* histidine biosynthesis and the associated genes? Phosphoribosylaminoimidazole-succinocarboxamide synthase (SAICAR synthetase; *purC*), Phosphoribosylformylglycinamidine cyclo-ligase (*purM*), bifunctional *purH*, Adenylosuccinate lyase (*purB*), and Phosphoribosylamine-glycine ligase (*purD*) were also noted as core essential genes of *N. gonorrhoeae* (Remmele *et al.*, 2014). Perhaps the relationship between histidine and purine biosynthesis could hint at driving factors to *N. gonorrhoeae* histidine gene retention.

Helicobacter pylori (*H. pylori*) is a histidine auxotroph with a heavily streamlined genome. As a result, *H. pylori* does not contain functional histidine biosynthesis genes, and additionally lacks a number of purine synthesis genes. *H. pylori* utilises the purine salvage pathway from IMP onward for purine requirements requiring *purA* (Bubić *et al.*, 2018). The retention of *purA* in such a reduced genome supports the idea that purine salvage pathways play a pivotal role in cell processes.

2.3.5 Does a single ATP-PRT isoform offer a significant benefit?

It is impossible to determine the advantages of one isoform over the other through literature. However, the two isozymes *G. sulfurreducens* and *D. acetoxidans,* offer a point of study. Significant work has been done in *G. sulfurreducens* by Aklujkar (2011) to determine if either isoform is redundant. Redundancies can be used to elucidate possible ancestral forms – or identify selection pressures for a particular regulatory domain. Both ATP-PRT isoforms are functional in *G. sulfurreducens,*

however, *hisG*$_S$ is more essential to growth than *hisG*$_L$. With the knockout of either isoform, the other would upregulate in response, signifying both isoforms are simultaneously functional and appropriately regulated within the organism (Aklujkar, 2011).

In the case of isozymes, if there was an evolutionary advantage to retaining both isoforms, such as surviving under nitrogen fixing conditions you could expect to see more ATP-PRT isozymes across bacterial species. We could assume in *G. sulfurreducens* and *D. acetoxidans* that the acquisition of the additional *hisG* is recent, and thus the redundant form has not yet been lost in genome reduction. For other species the lack of both forms supports the idea that *hisG* and the histidine biosynthetic pathway is ancient, and as a result regulatory mechanisms have had time to arise, spread, and reduce.

2.3.5.1 Is HisZ a transcriptional regulator?

Delorme *et al.* (1999) described T boxes that are frequently found in genes for aminoacyl-tRNA synthetases in bacteria. T boxes are signature sequences that act as positive regulators in the presence of uncharged tRNA. The presence of the specific amino acid alters the charge on tRNA creating a regulator. Uncharged, and therefore unregulated, tRNA is able to bind with leader sequences at the T box for regulation. It would be advisable to see whether histidyl-tRNA synthetase contain this T box, and thus would be able to function as transcriptional regulators of the histidine biosynthesis. It is possible the use of pre-transcriptional regulation allows energy conservation, and thus offers a substantial benefit to organisms with this mode of regulation.

2.4 Future experiments and recommendations

Constructing phylogenetic trees from a diverse bacterial group with low sequence similarity is challenging, and thus, inference based the phylogenetic trees is difficult. Creating a larger phylogenetic tree with a broader number of well annotated sequences would hopefully provide more information to create phylogenetic trees with more meaningful topology and use polytomous branching. Concatenating both HisG$_S$ and HisZ sequences in one MSA and PhyML tree may also assist in

observing where each gene arose and its common ancestry. Major substitution matrixes were tested during this research, but more could be tested to provide better estimations for reasonable mutation events, along with molecular clocks. Advances in the ease and accuracy of all these steps, will aide this extended analysis in the future.

There were multiple occurrences of histidine and purine biosynthesis crossovers discovered during this research. Sequentially knocking out each *pur*-gene and assessing the effect on histidine biosynthesis and *vice versa* would highlight any reliance of gene products on the other biosynthetic pathway.

AICAR was recognised as a competitive inhibitor in *E. coli* ATP-PRT; does this competitive inhibition occur in other species, and does this inhibition also occur in both isoforms in *in vitro* kinetic assays? If so, significantly lowered histidine production with a *purH* knockdown could be expected. Purine biosynthesis is tightly linked into histidine biosynthesis via substrate regeneration, and liberation of a possibly competitive compound (AICAR). This link is supported by *purA/hisZ* gene synteny, and other histidine and purine operon overlap in *Pelotomaculum thermopropionicum* and *D. audaxviator*.

It would also be worth investigating the role *hisZ* and *hisS* could play as a transcriptional regulator. This would help elucidate whether they are regulating the histidine or purine biosynthesis, or if they have been co-adopted throughout time.

2.5 Conclusion

Across bacterial groups *hisG*$_S$ appears to be the predominant gene and isoform. However, with the high sequence similarity of *hisZ*, to ancestral tRNA-synthetase, it is difficult to accurately quantify the percentage of long vs short isoforms.

Based on observations of distribution of *hisG*$_{S/L}$, HisG C-terminus, and *hisZ* synteny of surrounding genes, and tree topology we suggest *hisG*$_L$ is the ancestral variant of the two ATP-PRT forms. We suggest the ancestral Long-form diverged into the three domains of life, suggested by the distribution of the C-terminus across all three domains, and diverged within the bacteria with the complete loss of

C-terminus in the Firmicutes. The phylum Firmicutes is the only distinct group missing HisG_C, the gene synteny of histidine operon within the Firmicutes is well conserved with *hisZ* preceding the short-form *hisG*. The gene order of other bacterial short-forms is not as well conserved, with a prominent location of Proteobacteria $hisG_S$ following *purA*, suggesting HGT events disordering the operons.

What drives the selection pressure for $hisG_S$ over $hisG_L$ is unknown, however, *hisZ* complexed with $hisG_S$ acts as an enhancer to catalysis. It would be of interest to quantify the strength of allosteric regulation of histidine between the two groups. Perhaps a less stringent regulatory method allows for the overproduction of histidine. Potentially there is an increased demand for histidine production, and thus this requirement is greater than the requirement for ATP and PRPP conservation. Contrarily, perhaps histidine can be scavenged easily from the environment and thus the pathway can and is being used for another metabolic process.

The tight link between L-His and purine synthesis opens the possibility that this gene arrangement is an interplay of two complimentary pathways functioning for substrate recovery, allowing for the overproduction of histidine. This observation uncovers potential for kinetic work with the inclusion of PurA to investigate any interaction between the gene products of *purA*, *hisZ*, and *hisG*. Genes (including *hisG* and *hisZ*) involved in histidine biosynthesis are ranked higher than other amino acid synthesis genes in pathway essentiality and grouped among other major cell processes and energy cycles, hinting the *de novo* histidine biosynthetic pathway may not primarily function for the end point of histidine generation, but actually fit into a larger metabolic interplay.

Chapter 3
Materials and Methods

3.1 General methods

Each protein used in this research was purified using a similar immobilised metal affinity chromatography (IMAC) and gel filtration protocol to isolate active protein. Protein specific chromatography details, including methodology and buffers can be found in sections 3.3.1, 3.3.2, and 3.3.3.

3.1.1 Cloning of recombinant protein

The expression constructs for *mtu*ATP-PRT$_{WT}$, and *Eco*PPase, were kindly gifted by the Parker Lab (Victoria, University of Wellington). The expression construct for rTEV protease was kindly gifted by Dr Ghader Bashiri (University of Auckland). Recombinant proteins for *mtu*ATP-PRT$_{D258A}$ and *mtu*ATP-PRT$_{D258N}$ were purchased cloned into pET28a through TWIST BioScience. Gifted and purchased plasmids were transformed into chemically competent *E. coli* BL21 (DE3) and DH5α stocks *via* heat shock. Transformed *E.coli* was streaked onto LB agar containing appropriate antibiotics and incubated at 37 °C overnight. A single colony from each transformation was selected, and used to inoculate a 5 mL LB culture containing appropriate antibiotic, and incubated at 37 °C overnight shaking at 180 rpm. The following day liquid culture was combined, in cryotubes, with sterile glycerol to a final concentration of 25 %, and stored at -80 °C as glycerol stocks to be used in future experiments.

3.1.2 Colony PCR

*mtu*ATP-PRT$_{WT}$ intermittently did not express, as a result colony PCR was periodically used to ensure the gene of interest remained within the plasmid. The issue of gene loss was further resolved by the addition of 1 % glycerol to all plates and liquid cultures containing *mtu*ATP-PRTs. Colony PCR (composition described in Table 3.1) used a single *mtu*ATP-PRT colony suspended in 10 µL MQ H_2O as

template, T7 Forward (TAATACGACTCACTATAGGG) and Reverse (GCTAGTTATTGCTCAGCGG) were used in the amplification of pET28a constructs. PCR cycling was performed following conditions in Table 3.2 in a thermal cycler (Bio-Rad Laboratories, USA). PCR products were viewed on 1 % TAE (40 mM TRIS-acetate, 20 mM EDTA) agarose gel containing 1 x thiazole orange. PCR product size was determined by comparison to Invitrogen 1 Kb Plus DNA ladder (Thermo Fisher Scientific, USA).

Table 3.1. Colony PCR composition (20 µL reaction).

Component	Volume (µL)	Final concentration
5x HotFire Pol blend	4	1x
10 µM T7 Forward primer	0.6	0.3 µM
10 µM T7 Reverse primer	0.6	0.3 µM
Template	1	-
MQ H_2O	13.8	-

Table 3.2 PCR cycling conditions for T7 colony PCR of *mtu*ATP-PRTs.

	Steps	Temperature (°C)	Length (minutes)
	Pre-denaturation	95	15:00
x 29	Denaturation	95	00:30
	Annealing	55	00:20
	Extending	72	00:30
	Final extension	72	07:00

3.2 Expression of recombinant proteins

3.2.1 Expression of rTEV protease

Frozen glycerol stocks were spread onto LB agar (containing 50 mg/L kanamycin) and incubated at 37 °C overnight. Single colonies were selected and inoculated into 20 mL LB broth containing 50 mg/L kanamycin, and grown overnight at 37 °C, shaking at 180rpm. A 1 L expression culture of TEV autoinduction medium (**Table 3.8**) containing 50 mg/L kanamycin was inoculated with the 20 mL starter culture

and grown at 28 °C for 20 hours. Cells were harvested by centrifugation at 4500 g for 20 minutes at 4 °C, and frozen at -80 °C until use.

3.2.2 Expression of *Eco*PPase

Frozen glycerol stocks were spread onto LB agar (containing 100 mg/L ampicillin) and incubated at 37 °C overnight. Single colonies were selected and inoculated into 20 mL LB broth containing 100 mg/L ampicillin, and grown overnight at 37 °C, shaking at 180 rpm. An expression culture of 1 L LB containing 100 mg/L ampicillin was inoculated with the 20 mL starter culture, and grown at 37 °C, shaking at 180 rpm until log phase was reached (OD_{600} 0.6-0.8). Protein expression was then induced by the addition of 0.75 mM IPTG, and temperature adjusted to 23 °C for an additional 20 hours. Cells were harvested by centrifugation at 4500 g for 20 minutes at 4 °C, and frozen at -80 °C until use.

3.2.3 Expression of *mtuATP-PRT*$_{WT}$, *mtuATP-PRT*$_{D258A}$, and *mtuATP-PRT*$_{D258N}$

1 µL of glycerol stock was inoculated into 40 mL LB broth containing 50 mg/ mL kanamycin and 1 % glucose, and grown overnight at 37 °C, shaking at 180 rpm. A 1 L ZY media culture (Studier, 2005) containing 50 mg/L kanamycin was inoculated with the 40 mL starter culture, and grown at 37 °C, shaking at 180 rpm for 24 hours. Cells were harvested by centrifugation at 4500 g for 20 minutes at 4 °C and used immediately.

3.2.4 Cell lysis

Cell pellets were suspended in 3x cell pellet volume of lysis or binding buffer (protein specific; see Tables 3.3-3.5) and defrosted over ice. *Eco*PPase was treated with the addition of a protease inhibitor prior to sonication (Roche cOmplete Mini Proteoase Inhibitor Cocktail). Resuspended cells were lysed by sonication (QSonica, 12-amp, 3-minute processing time, 1 second on, 1 second off) in a Bioscience cool block. Cell debris was removed by centrifugation 20,000 g, for 20 minutes at 4 °C to isolate protein containing supernatant.

3.3 Purification of active recombinant protein

3.3.1 Purification of TEV protease

Following cell lysis (in binding buffer), protein containing supernatant was filtered to 0.2 μm (Minisart syringe filters; Sartorious AG, Germany), and loaded onto 5 mL IMAC column (HiTrap HP; GE Healthcare, UK) pre-equilibrated with binding buffer (**Table 3.3**). Weakly bound non-target proteins were eluted from the column with a solution comprising of 4 % buffer B: 96 % buffer A at a flow rate of 1 $ml.min^{-1}$. Protein isolation from the IMAC column was achieved using an ÄKTA Basic, Prime, or Purifier FPLC system (GE Healthcare, Sweden) using an isocratic gradient from 4-100 % elution buffer over a 70 mL gradient.

Protein containing fractions were pooled and placed in 6-8 kD MWCO dialysis tubing (Spectrum™ Spectra/Por™; Fisher Scientific, USA) and dialysed overnight in dialysis buffer (**Table 3.3**) at 4 °C. Dialysed protein was collected, quantified using NanoDrop (Thermo Fisher Scientific, USA), and sterile glycerol added to a final concentration of 50 %. This solution was aliquoted in 1 mL volumes, snap frozen in liquid N_2 and stored at -80 °C until use.

Table 3.3 Buffer compositions for IMAC and dialysis of TEV protease.

Buffer	Composition
Binding/lysis buffer	50 mM TRIS, 500 mM NaCl, 1 mM β-ME, 10 % glycerol, pH 7.5
IMAC buffer A	50 mM TRIS, 500 mM NaCl, 1 mM β-ME, 10 mM imidazole, pH 7.5
IMAC buffer B	50 mM TRIS, 500 mM NaCl, 1 mM β-ME, 1M imidazole, pH 7.5
Dialysis buffer	20 mM TRIS, 150 mM NaCl, 0.5 mM β-ME, pH 7.5

3.3.2 Purification of *Eco*PPase

Following cell lysis, protein containing supernatant was filtered to 0.2 μm (Minisart syringe filters; Sartorious AG, Germany), and loaded onto 5 mL IMAC column (HiTrap HP; GE Healthcare, UK) pre-equilibrated with binding buffer (**Table 3.4**)

Weakly bound non-target proteins were eluted from the column with a solution comprising of 4 % elution buffer: 96 % lysis buffer at a flow rate of 1 ml.min^{-1}. Protein isolation from the IMAC column was achieved using an ÄKTA Basic, Prime, or Purifier FPLC system (GE Healthcare, Sweden) using an isocratic gradient from 4-100 % elution buffer over a 50 mL gradient.

The major protein elution peak was identified following the 280 nm trace, and concentrated to <1 mL volume in a 20 mL Vivaspin concentrator (10 kDa molecular weight cut off; Sartorius AG, Germany) at 2600 g, at 4 °C. Concentrated protein was filtered to 0.2 µm and injected onto ENrich™ SEC 650 10 x 300 column (Bio-Rad Laboratories, USA) pre-equilibrated with size exclusion buffer (**Table 3.4**). Protein was separated and eluted with size exclusion buffer at a flow rate of 0.5 ml.min^{-1} and collected in 0.5 mL aliquots. Fractions containing protein were identified by following the 280 nm wavelength trace. Fractions of interest were combined and aliquoted in 50 µL volumes, snap frozen in liquid N_2 and stored at -80 °C until use.

Table 3.4 Buffer compositions for IMAC, and size exclusion chromatography of *Eco*PPase.

Buffer	**Composition**
Binding/IMAC lysis buffer	50 mM $NaPO_4$, 500 mM NaCl, 5 mM $MgCl_2$, 20 mM imidazole, pH 8.0
IMAC elution buffer	50 mM $NaPO_4$, 500 mM NaCl, 5 mM $MgCl_2$, 500 mM imidazole, pH 8.0
Size exclusion buffer	50 mM TRIS, 100 mM NaCl, 5 mM $MgCl_2$, pH 8.0

3.3.3 Purification of *mtuATP-PRT*$_{WT}$, *mtuATP-PRT*$_{D258A}$, and *mtuATP-PRT*$_{D258N}$

Cells were lysed by sonication (Section 3.2.4). Collected supernatant was filtered to 1.2 µm (Minisart syringe filters; Sartorious AG, Germany), and loaded onto 5 mL IMAC column (HiTrap TALON crude; GE Healthcare, UK) equilibrated with IMAC binding buffer (**Table 3.5**). Weakly bound non-target proteins were eluted from the column with a solution comprising of 100 % buffer A at a flow rate of 1 ml.min^{-1}. Protein isolation from the IMAC column was achieved using an ÄKTA

Basic, Prime, or Purifier FPLC system (GE Healthcare, Sweden) using an isocratic gradient from 0-100 % buffer B over a 35 mL gradient.

The fractions of the major peak (250 mM imidazole) as identified using the 280 nm absorbance were collected and pooled. Protein concentration was measured via NanoDrop (Thermo Fisher Scientific, USA). Pooled protein was diluted 4-fold in size exclusion buffer (**Table 3.5**) to reduce NaCl to approximately 100 mM for optimised TEV activity, before the addition of 1 mg TEV protease. Protein was incubated at 37 °C for 1 hour followed by 4 °C overnight for the removal of poly-histidine tags.

After TEV protease digestion, protein mixture was concentrated to 5 mL in a 20 mL Vivaspin concentrator (10 kDa molecular weight cut off; Sartorius AG, Germany) at 2600 g, at 4 °C. Concentrated protein was removed from the concentrator and centrifuged for 10 minutes at 4500 g for sedimentation of precipitated protein and dust before injection onto HiLoad 16/600 Superdex 200 column (GE Healthcare Life Science, UK) pre-equilibrated with size exclusion buffer (**Table 3.5**). Protein was eluted at 0.5 mL/minute as 2 mL fractions. Fractions containing protein were identified by following 280 nm wavelength trace. Purified protein was aliquoted as 500 µL or 1 mL volumes, snap frozen in liquid N_2, and stored at -80 °C. Protein was thawed over ice before onetime use.

Table 3.5 Buffer compositions for IMAC and size exclusion of *mtuATP-PRT*$_{WT}$, *mtuATP-PRT*$_{D258A}$, and *mtuATP-PRT*$_{D258N}$

Buffer	**Composition**
IMAC binding/lysis buffer	50 mM $NaPO_4$, 500 mM NaCl, pH 7.8
IMAC buffer A	50 mM $NaPO_4$, 500 mM NaCl, 5 mM imidazole, pH 7.8
IMAC buffer B	50 mM $NaPO_4$, 500 mM NaCl, 500 mM imidazole, pH 7.8
Size exclusion buffer	50 mM TRIS, 100 mM NaCl, 5 mM $MgCl_2$, pH 8.0

3.4 Identification of target protein

3.4.1 DNA sequencing of plasmid

To confirm the sequence identity of each recombinant protein the plasmid was purified and sequenced between T7 promotor sequences. 5 mL LB and antibiotic mixture was inoculated with the glycerol stock for each protein and grown overnight at 37 °C, 180 rpm. Overnight cultures were centrifuged at 4500 g, for 10 minutes, at 4 °C, and plasmid extracted using the QIAprep Spin Miniprep Kit (Qiagen, Germany). Purified plasmid was diluted to 250-625 ng with 4 pmol of each T7 forward, and T7 reverse and sequenced using Sanger sequencing (Massey Genome Service, New Zealand). Returned sequences were trimmed, aligned, and mapped to reference sequence using Geneious Prime (Geneious Prime 2019.2.1 (https://www.geneious.com), Biomatters Ltd, New Zealand) to ensure target DNA sequence was accurate and free of mutations before expression.

3.4.2 Analytical size exclusion

The analytical size exclusion column (ENrich™ SEC 650 10 x 300 column, Bio-Rad Laboratories, USA) was first equilibrated with the protein specific size exclusion buffer, then calibrated with 1 $mg.ml^{-1}$ Blue Dextran and Gel filtration standard (Bio-Rad Laboratories, USA). Blue Dextran and Gel filtration standards were used to calculate the void volume and a standard curve of protein size respectively. 250 µL of purified recombinant protein was run through the analytical size exclusion column, and the elution volume of protein was used against the standard curve to determine protein size and oligomeric arrangement.

3.4.3 Polyacrylamide Gel Electrophoresis (PAGE)

PAGE gels were used routinely to assess purity and quality of recombinant expressed proteins. Gels were cast 5 at a time, and stored at 4 °C. All gels were made as 12 % native gels (**Table 3.6**), however, samples were prepared in a 4 x SDS loading buffer (**Table 3.7**), and run in 1x TG-SDS (**Table 3.7**). Protein bands were stained with Fairbanks stain and de-stained (**Table 3.7**). Denatured protein band product size was determined by comparison to Precision Plus Protein standards (Bio-Rad Laboratories, USA).

Table 3.6 Components of 12 % Polyacrylamide gels (Native) to cast 5 gels.

Component	Resolving gel	Stacking layer
MQ H_2O	10.35 mL	8.63 mL
30 % acrylamide	12 mL	2.125 mL
TRIS	7.5 mL (1.5 M, pH 8.8)	1.6 mL (1 M, pH 6.8)
10 % APS	150 µL	63 µL
TEMED	15 µL	6.3 µL

Table 3.7 Preparation, running, and visualisation buffers used in PAGE.

	Components	Methods
4 x SDS buffer	250 mM TRIS (pH 6.8), 20 % (v/v) glycerol, 4 % (w/v) SDS, 10 % (v/v) β-ME, 0.025 % (w/v) bromophenol blue	Dilute four-fold in protein sample, incubate at 95 °C for five minutes
1 x TG-SDS	25 mM Tris, 250 mM glycine, 0.1 % SDS	
Fairbanks stain	0.05% coomassie blue (R250), 25 % (v/v) isopropanol, 10 % (v/v) acetic acid	Submerge PAGE in stain, microwave for 30 seconds, shake for five minutes at room temperature
De-stain	10 % (v/v) acetic acid	Submerge PAGE in de-stain, microwave for 30 seconds, shake for 20 minutes at room temperature with a tissue to absorb dye

3.5 Stock materials

3.5.1 Expression Medium

Table 3.8 Media components used in the expression of recombinant proteins TEV protease, EcoPPase, and *mtu*ATP-PRT.

Media	Composition
LB Broth	10 g Peptone, 5 g yeast extract, 10 g NaCl
TEV expression media	12 g Tryptone, 24 g yeast extract, 8 mL glycerol, 5 g lactose, 0.15 g glucose, 2.314 g KH_2PO_4, 12.542 g K_2HPO_4, 0.24 g $MgSO_4$, 3.7 g aspartic acid, 1.2 g NaOH
ZY Media	10 g Peptone, 5 g yeast extract, 20 mL 50x M stock, 20 mL 50x 5052 stock, 1 mL 1 M $MgSO_4$, 200 µL 1000x trace metal stock
50x M	1.25 M Na_2HPO_4, 1.25 M KH_2PO_4, 2.5 M NH_4Cl, 0.25 M Na_2SO_4
50x 5052	25 % (w/v) Glycerol, 2.5 % (w/v) Glucose, and 10 % (w/v) Lactose
1000x Trace metal mix	0.1 M $FeCl_3$, 1 M $CaCl_2$, 1 M $MnCl_2$, 1 M $ZnSO_4$, 0.2 M $CoCl_2$, 0.1 M $CuCl_2$, 0.2 M $NiCl_2$, 0.1 M $NaMoO_4$, 0.1 M Na_2SeO_3, and 0.1 M H_3BO_3

3.5.2 Antibiotic stocks

Antibiotics used in growth media were diluted from 1000x concentrated stocks that were prepared in bulk and stored at -20 °C until use. Stocks were solubilised in MQH_2O (unless otherwise stated) at concentrations of; ampicillin (100 mg/mL), and kanamycin (50 mg/mL).

3.6 Differential scanning fluorimetry based thermal shift assay

Differential scanning fluorometry using SYPRO Orange (Life Technologies, USA) were used to assess the thermal stability of *mtu*ATP-PRT and the shift to thermal stability upon allosteric ligand binding (L-His and TIH) (Lo *et al.*, 2004). Protein

and SYPRO Orange concentrations were held at 1 mg/mL and 50 X respectively. Ligands were serially diluted to ensure each reaction had the same volume added. L-His varied from 0, 10, 20, 30, 40, and 50 µM. TIH was varied from 0, 0.25, 0.5, 1, 4, and 8 mM. Each reaction was run in triplicate along with blank reactions containing ligand and no protein in a Rotor-Gene RT-PCR machine (Corbett Life Science, Australia). A temperature gradient of 25-99 °C was used, with continuous monitoring of fluorescence (excitation 470, emission 550 nm).

Data was analysed by taking the first derivative of the curve for each run, the inflection point of this data corresponded to the melting temperature, which were averaged and plotted against ligand concentration in GraphPad Prism 9.0 (GraphPad Software, USA).

3.7 Kinetic assays

3.7.1 Substrate preparation

Due to the complexity and sensitivity of *mtu*ATP-PRT assays, substrates and ligands were prepared in advance and stored at -20 °C, and thawed over ice, before onetime use. PRPP appeared to phosphorylate overtime, changing concentration and was hydroscopic making it difficult to weigh at small amounts. Therefore, entire bottles would be solubilised at a time and stored at -20 °C so any phosphorylation would remain consistent across the batch. Substrates and ligands were solubilised in assay buffer (**Table 3.10**) to retain balance of $MgCl_2$ and KCl co-factors.

Table 3.9 Concentration of substrates and ligands prepared for *mtu*ATP-PRT assays. Substrates and ligands were suspended in assay buffer and stored at -20 °C.

Component	Concentration
PRPP (Sigma-Aldrich, USA; Cayman Chemical, USA)	100 mM
ATP (Sigma-Aldrich, USA)	250 mM
L-His (Sigma-Aldrich, USA)	100 mM
TIH (Sigma-Aldrich, USA)	600 mM

3.7.2 Assay buffers

Multiple buffers were used during the optimisation of *mtu*ATP-PRT assays. Publication of ATP-PRT's use a TRIS buffer (Mittelstädt *et al.*, 2016; Pisco *et al.*, 2017), however, due to the extensive shift of pH across temperature, TRIS was inappropriate in the scope of temperature data. Alternative buffers were tested for optimal activity, however, required adjustments to pH and co-factor concentrations due to some ability to bind free metal ions. Buffers were selected from the Good's Buffer table, KCl and $MgCl_2$ concentration was varied.

Table 3.10 Buffer composition for *mtu*ATP-PRT assays.

Buffer	**Composition**
TRIS buffered (Mittelstädt *et al.*, 2016)	50 mM TRIS, 100 mM NaCl, 50 mM KCl, 10 mM $MgCl_2$, pH 8.5
TRIS buffered (Pisco *et al.*, 2017)	20 mM TRIS, 200 mM KCl, 7 mM $MgCl_2$, pH 8.5
Phosphate buffered	50 mM Phosphate, 100 mM NaCl, 50 mM KCl, 10 mM $MgCl_2$, pH 7.4
Phosphate buffered	50 mM Phosphate, 100 mM NaCl, 200 mM KCl, 10 mM $MgCl_2$, pH 7.4
Phosphate buffered	50 mM Phosphate, 100 mM NaCl, 200 mM KCl, 20 mM $MgCl_2$, pH 7.4
Phosphate buffered	50 mM Phosphate, 100 mM NaCl, 200 mM KCl, 60 mM $MgCl_2$, pH 7.4
Boric Acid buffered	50 mM Boric acid, 100 mM NaCl, 50 mM KCl, 10 mM $MgCl_2$, pH 9.0
Boric Acid buffered	50 mM Boric acid, 100 mM NaCl, 200 mM KCl, 10 mM $MgCl_2$, pH 9.0

3.7.3 Kinetic Assays

3.7.3.1 Continuous assays

Continuous assays were performed in a 200 or 500 µL quartz cuvette in a Heλios γ spectrophotometer (ThermoFisher, USA) temperature controlled by a single cell peltier (Thermo Scientific, USA) or 160 µL cell in a TgK Scientific Stopped-flow (Hi-Tech Scientific, UK) temperature controlled by a circulating water bath. Cuvette temperature was confirmed via thermocouple. Assay protocols (typically 450 nM *mtu*ATP-PRT, 500 nM *Eco*PPase, and varying concentration of substrates) based on (Ames, Martin, & Garry, 1961) were initiated with PRPP. Progression was continuously read every 0.125 seconds at 290 nm in Vision32 software (version 1.25) for Heλios γ, or read every 0.06 seconds at 290 nm in Kinetic Studio (version 4.0.12) for the Stopped-flow. Heλios γ was used during optimisation of substrate and buffer optimisation due to the low sample volume required, and Stopped-flow was used for final data collection of T_{opt}, IC_{50}, and Michaelis-Menten plots.

Michaelis-Menten plots were collected for both substrates for each enzyme depending on substrate availability. Due to shipping delays stocks of PRPP were very low as is reflected in the breadth of some Michaelis-Menten plots and concentrations used in final assay collection. Michaelis-Menten plots were completed with 450 nM *mtu*ATP-PRT, 500 nM *Eco*PPase, varying concentrations of PRPP (ATP held at 3 mM final) or varying concentrations of ATP (PRPP held at 1.5 mM). Substrates of interest were serially diluted, and collected from lowest to highest concentration at described temperatures. IC_{50} data for L-His and activation data for TIH were also collected using serial dilution methods to guide appropriate concentrations of inhibitors and activators that allowed activity to remain visible, but modulation to be observed.

3.7.3.2 Temperature Optimums

Michaelis–Menten results were used to optimise conditions for T_{opt}'s where reasonable. Temperature optimums were collected at saturating substrate (where possible), with allosteric ligand concentrations based on IC_{50} results. All temperature curves were collected via continuous spectrophotometric assay on the TgK Scientific Stopped-flow with a T-cell for extended temperature range.

Orientation and composition of the two drive syringes was kept consistent; enzyme, ligand (if applicable), *Eco*PPase, and ATP were combined in the left syringe with PRPP alone in the right syringe. The reaction would then be initiated in the cell. Temperature was regulated via circulating water bath connected to the observation cell of the sample handling unit. Temperature curves were collected over a 45 °C range from 15 °C to 60 °C by increasing temperature approximately 3 ° for each collection. Data points for each temperature were collected in triplicate with 5 'dummy shots' of 0.2 seconds length between each data collection of 45-60 seconds to ensure the clearance of PR-ATP product build-up within the observation cell and adequate heating time of the reaction cell, respectively.

3.7.3.3 Analysis

Rate data was analysed by linear regression from linear data (approximately 5 seconds) using Vision32 or Kinetic studio. Data analysis and fitting was performed in GraphPad Prism 9.0 (GraphPad Software, USA).

3.8 Crystallography

3.8.1 Screening and crystal conditions

Hanging drop vapour diffusion crystal screens were generated following previously published crystallisation methods of 0.1 M MES (pH 6.1-6.7) and $MgSO_4$ (0.6-1.6 M), or 2.5 % isopropanol and $(NH_4)_2SO_4$ (0.6-1.2 M) (Cho *et al.*, 2003; Pisco *et al.*, 2017). Protein and ligand concentration were varied in the generation of well-formed crystals. Crystals were grown at 18 °C, co-crystallised with ligand L-His (50-200 µM), or TIH (1-5.5 mM).

Crystal morphology was limited to large triangular prism or large cubes. All were cryoprotected in 20 % glycerol rapidly submerged in liquid nitrogen for data collection.

3.8.2 Data collection and processing

Diffraction data was collected by X-ray diffraction. Data was collected using the MX2 beamline (Australian Synchrotron, Melbourne, Australia) with an EIGER X 16M pixel detector (Dectris Ltd, Switzerland). Diffraction data was collected 360 ° over 36 seconds, with the detector between 250-350 mm, and attenuation varying between 60 – 90 %. Collected data was indexed, integrated and scaled using XDS (Kabsch, 2010), and space group determined with Pointless (Evans, 2011).

3.8.3 Structure solution

Datasets were scaled using AIMLESS (Evans & Murshudov, 2013) (CCP4 suite version 7.1.013) (Winn *et al.*, 2011) until R-merge values were between 0.8 - 1.3 in the outer shell. 5 % of data was reserved for R-free. Matthews Cell Contents Analysis (Matthews, 1968; Kantardjieff & Rupp, 2003) with molecular weight set to 32330 Daltons (CCP4 suite version 7.1.013) was performed on each dataset to confirm the contents of the asymmetric unit.

Scaled data was imported to Phenix (version 1.18.2-3874) (Liebschner *et al.*, 2019) and Phaser-MR (full-featured) (McCoy *et al.*, 2007) was used for molecular replacement model building. With waters, ligands, and molecules removed PDB: 1NH7, 5LHU, and 5LHT (Cho *et al.*, 2003; Pisco *et al.*, 2017) were used as models for molecular replacement of APO, Histidine bound, and TIH bound structures, respectively.

Following molecular replacement structure and difference map files were bought into Coot (version 0.9.41) (Emsley *et al.*, 2010), for manual refinement. $2F_O$-F_C electron density map was contoured to 1 σ for the duration of refinement. Multiple rounds of refinement were performed through Refmac5 (CCP4 suite version 7.1.013) (Murshudov *et al.*, 1997) and phenix.refine (Afonine *et al.*, 2012) until the crystal model was satisfactory and R-work and R-free values plateaued around 0.2 and 0.25 respectively. Refmac5, Coot, and Phaser-MR (Phenix suite) were used as tools for the addition of water molecules. Structures were submitted to PDB validate to assess structure quality and to highlight structural issues that were resolved through further rounds of refinement and editing in Coot.

3.9 Structural analysis

PyMOL (Schrödinger, Inc. version 2.3.3) was used for the visualisation of structures. Structures were compared by eye using alignment tools in PyMOL. The hexameric structure of ATP-PRT was generated through symmetry mates in PyMOL, and each chain was merged into a single object. B-factors of the six solved structures and 1NH7, 5LHU, and 5LHT were calculated in BAVERAGE (Dodson, 1991) in CCP4. RMSD between structures was calculated in PyMOL by sequentially comparing each of the 9 structures to each other. *mtu*ATP-PRT$_{WT}$ structures PDB: 5LHU and 5LHT were compared using Phenix Structure comparison to identify rotamers, B-factors, and Ramachandran outliers to guide structure comparison. Structure topology was designed in PDBsum (Laskowski *et al.*, 1997).

3.10 Hydrogen bond analysis

Hexamers of each structure generated through symmetry mates were imported to Chimera X (version 1.2.5) (Pettersen *et al.*, 2004). H-bonds were calculated and written out to a text file. Data was processed, identifying and extracting novel bonds existing for each mutant in different ligand bound states using a Python script developed by Carlin Hamill (University of Waikato) for this purpose (Appendix 2). Novel bonds were sorted by chain interactions, and intra and inter-chain interactions involving chain A were exported into PyMOL with existing structures and labelled.

3.11 Molecular Dynamics

Crystal structures were prepared for molecular dynamics by filling in all backbones of the structure that were missing from density. Ligands and waters were left in the structure. The hexameric structure of ATP-PRT was generated through symmetry mates in PyMOL, and each chain was merged into a single object. Molecular

dynamic experiments were run by Dr Wanting Jiao (Victoria University of Wellington) using the Desmond Suite (Schrödinger) (Bowers *et al.*, 2006).

3.12 Isothermal Titration Calorimetry

Purified *mtu*ATP-PRT$_{WT}$ was buffer exchanged into kinetic assay buffer (50 mM Boric acid, 100 mM NaCl, 50 mM KCl, 10 mM $MgCl_2$, pH 9.0). Allosteric ligand L-His was solubilised in the identical assay buffer. Isothermal titration calorimetry (ITC) data was kindly collected by Dr Gerd Mittelstädt (Victoria University of Wellington) due to travel constraints.

Dissociation constants at all three temperatures were determined using an Affinity-ITC microcalorimeter (TA Instruments). Solutions were filtered and degassed under vacuum and the protein concentration was measured by UV absorption prior to the experiments (concentration of *mtu*ATP-PRT$_{WT}$ was held at 185 µM). The cell was washed multiple times with degassed water followed by binding buffer and degassed binding buffer.

For all experiments the cell was loaded with the protein and the syringe contained the ligand (titrant). Titrations were performed with 26 injections (one 0.3 uL injection followed by 25 x 1 uL injections) of ligand. Heat of dilution was measured separately and subtracted from the integrated data before data-fitting in NanoAnalyze (version 3.12.0). The initial point (0.3 uL injection) was routinely deleted to allow for diffusion of ligand across the needle tip during the equilibration period and the binding model used was chosen to be individual. Confidence levels for error determination were set to 95 %.

Chapter 4
Kinetic characterisation of *mtu*ATP-PRT under allosteric modulation

4.1 Introduction

Allosteric inhibition of ATP-PRT by L-Histidine has been thoroughly reported in the literature (Pedreño *et al.*, 2012; Zhang *et al.*, 2012; Kulis-Horn *et al.*, 2015; Mittelstädt *et al.*, 2016; Pisco *et al.*, 2017). Studies of ATP-PRT regulation have been supplemented with crystallographic data to illustrate the significant movement of the regulatory domain forming a tensed structure less able to catalyse the reaction (Cho *et al.*, 2003; Mittelstädt *et al.*, 2016; Pisco *et al.*, 2017; Jiao *et al.*, 2019). Interestingly, using crystallographic data, allosteric activation by TIH implicates the same tensed conformation, yielding increased catalysis in *mtu*ATP-PRT. Converging structural changes with diverging catalytic rates upon allosteric modulation suggests tensed and relaxed protein states alone may not control allosteric regulation and that protein crystallography is unable to capture the dynamic movement of the enzyme under allosteric modulation. Traditional steady-state kinetic investigations of allosteric enzymes measure catalytic rate at a fixed temperature. This chapter applies temperature as an additional variable in kinetic assays to investigate the potential of fluctuating dynamic processes associated with allosteric regulation and investigate the effect of selective D258 mutation on the rate and allosteric response to both L-His (inhibitor) and TIH (activator).

4.2 Results

4.2.1 Expression and purification of *mtu*ATP-PRT and its mutants

Previous MD simulations by Jiao *et al.* (2019) probed for residues involved in the signal of allosteric inhibition by L-His in *cje*ATP-PRT. MD identified a number of candidate residues, one candidate, *cje*ATP-PRT$_{E271Q}$ was used to guide mutation in *mtu*ATP-PRT. Structural comparison between *mtu*ATP-PRT and *cje*ATP-PRT identified residue D258 occupying the same position as E271 from *cje*ATP-PRT.

*mtu*ATP-PRT$_{D258A}$ and *mtu*ATP-PRT$_{D258N}$ were generated to identify if alteration to this hinge region altered signal transduction of either allosteric ligand.

Recombinant protein expression in *E. coli* (Section 3.2) was used to produce soluble, active protein used in thermal melts, and continuous assays for all variants. *mtu*ATP-PRT$_{WT}$ and the mutant enzymes *mtu*ATP-PRT$_{D258A}$ and *mtu*ATP-PRT$_{D258N}$ were expressed in *E. coli* BL21 (DE3) under standard conditions (auto induction media, 37 °C, 24 hours). A two-step purification *via* IMAC followed by TEV digest, and gel filtration chromatography produced consistent yields of pure protein suitable for storage, kinetic assays, and crystallography for *mtu*ATP-PRT$_{WT}$ (**Figure 4.1**) and *mtu*ATP-PRT$_{\Delta D258}$ variants (**Figure 4.2**).

TEV digest was performed to align with APT-PRT work by Mittelstädt *et al.* (2016). Complete digestion of hexameric *mtu*ATP-PRT$_{WT}$ proved difficult, and this was also acknowledged during the purification of *cje*ATP-PRT. It appears some His-tags of the hexamer may be inaccessible to TEV protease, and as a result mixed populations of tagged and un-tagged chains were visible. These populations were not distinguished through preparative gel filtration chromatography, or analytical size exclusion chromatography. Reverse His-trap protocols of gel filtration purified protein over Ni beads also resulted in no change to tagged and untagged populations visible on SDS-PAGE (**Figure 4.2**). Gel filtration chromatography purified enzyme was consistently active in kinetic assays, and crystallised despite a mixed population of tagged and un-tagged monomers. To mitigate any potential variation with mixed populations of tagged and un-tagged monomers each *mtu*ATP-PRT variant was purified in large scale to ensure each set of assays was collected using the same purification batch.

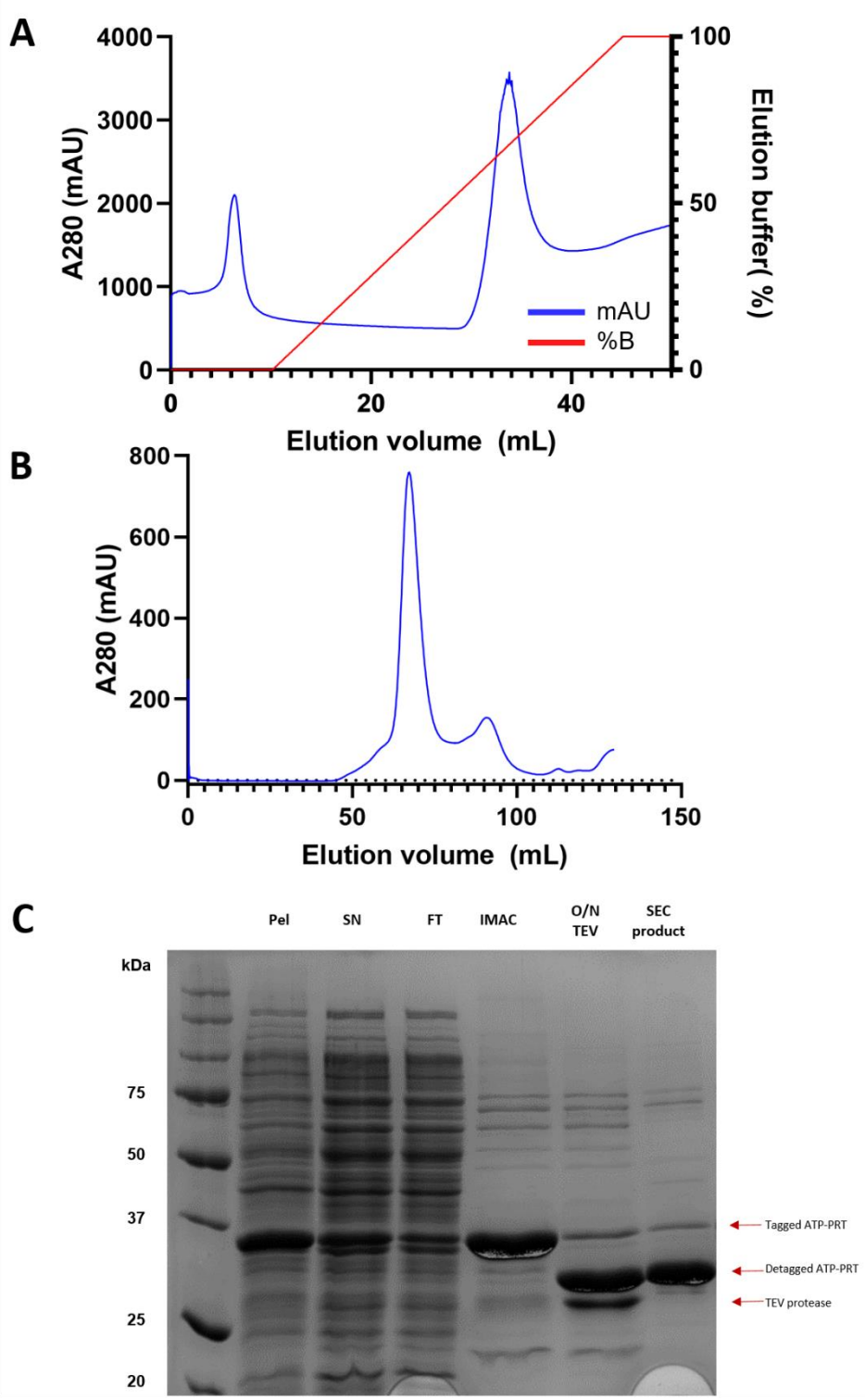

Figure 4.1. Purification process of *mtu*ATP-PRT$_{WT}$ comprising IMAC, gel filtration, and visualisation on SDS-PAGE (Section 3.4.3). A. IMAC chromatogram of protein elution on isocratic gradient with increasing imidazole (0 – 0.5 M). The major *mtu*ATP-PRT$_{WT}$ peak is taken for downstream purification process. B. Gel filtration chromatogram of protein (S200 16/60 gel filtration column). The major peak (65 mL) represents a population of purified *mtu*ATP-PRT. The peak at 90 mL represents TEV. C. Denaturing PAGE (12 %) of *mtu*ATP-PRT with samples from major IMAC peak (A), TEV digest efficiency, and gel filtration purified *mtu*ATP-PRT$_{WT}$. His-tagged *mtu*ATP-PRT is at 37 kDa, de-tagged *mtu*ATP-PRT at 32 kDa, TEV protease 28 kDa, and His-tags 3 kDa (non-resolvable on 12% SDS-PAGE).

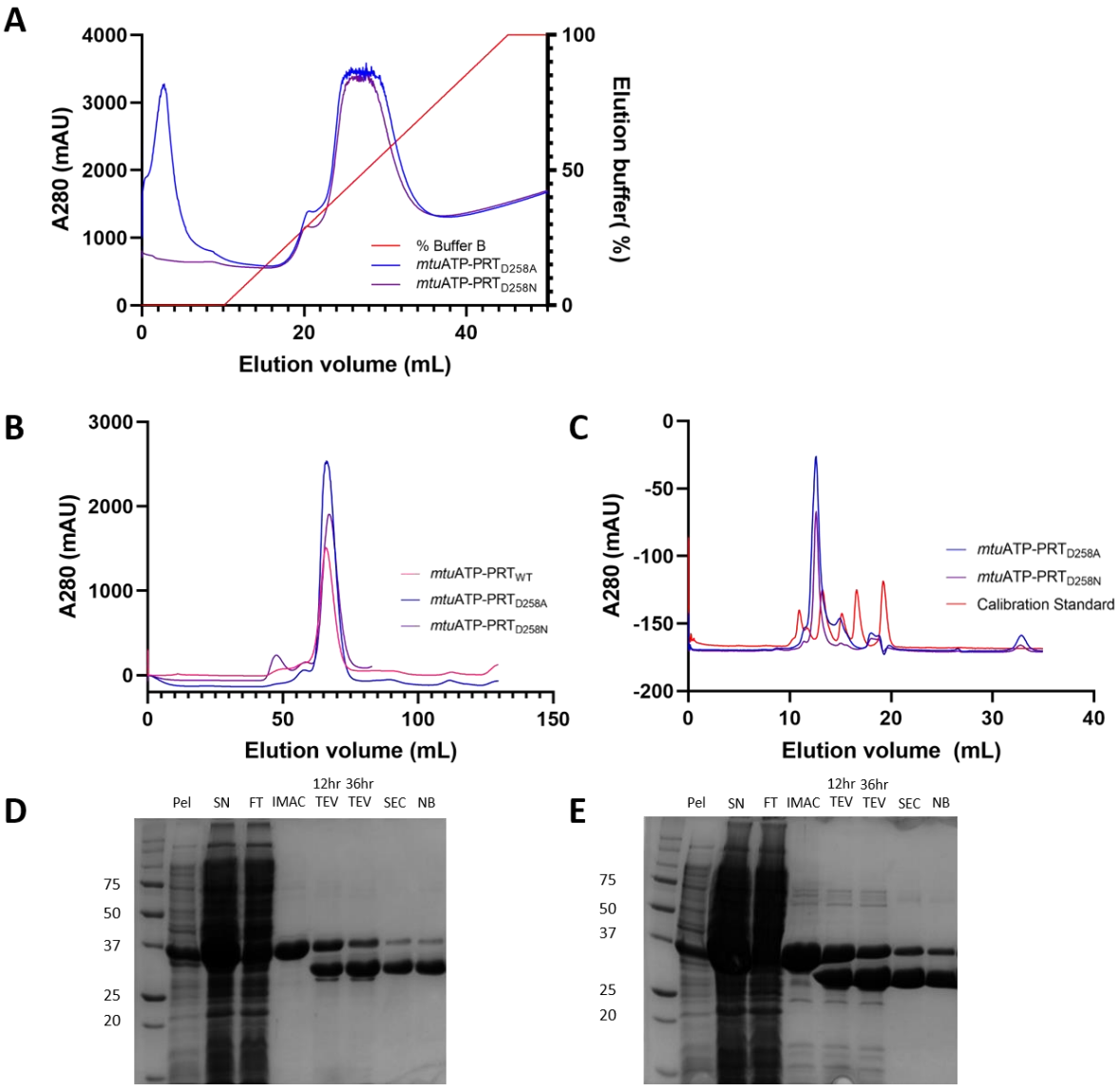

Figure 4.2. Purification process of *mtu*ATP-PRT$_{D258A}$ and *mtu*ATP-PRT$_{D258N.}$ **A**. IMAC chromatogram of *mtu*ATP-PRT$_{\Delta 258}$ proteins eluted on isocratic gradient with increasing imidazole (0 – 0.5 M) (Section 3.3.3). **B**. Preparative gel filtration (GE S200 16/60) chromatogram of *mtu*ATP-PRT variants in reference to wildtype. **C**. Analytical gel filtration chromatogram of protein (Biorad EnRich650 gel filtration column) identifies *mtu*ATP-PRT variants at 272.8 kDa (indicative of hexameric arrangement). **D**. and **E**. Denaturing PAGE (12 %) of *mtu*ATP-PRT$_{D258A}$ (**D**) and *mtu*ATP-PRT$_{D258N}$ (**E**) IMAC products, TEV digest efficiency, size exclusion product, and enzyme remaining in solution followed by a wash over nickel beads (NB). For both **D** and **E** His-tagged *mtu*ATP-PRT is at sized at 37 kDa. His-tags could not be completely removed through TEV digest, however, the purified product from size exclusion did not appear to have any affinity to Nickel resin, showed good activity in assays, and crystallised well. These figures show *mtu*ATP-PRT$_{\Delta 258}$ variants express and purify to high concentrations similarly to wild type.

4.2.2 Thermal stability of *mtu*ATP-PRT

Differential scanning fluorimetry (DSF) (Section 3.6) using SYPRO orange was used to compare the melting temperature (T_m) of *mtu*ATP-PRT$_{WT}$ in allosteric

modulated states. The melt trace of both L-His and TIH bound *mtu*ATP-PRT$_{WT}$ is shown in **Figure 4.3** and the maximal T_m of each concentration of ligand is shown in **Figure 4.4**.

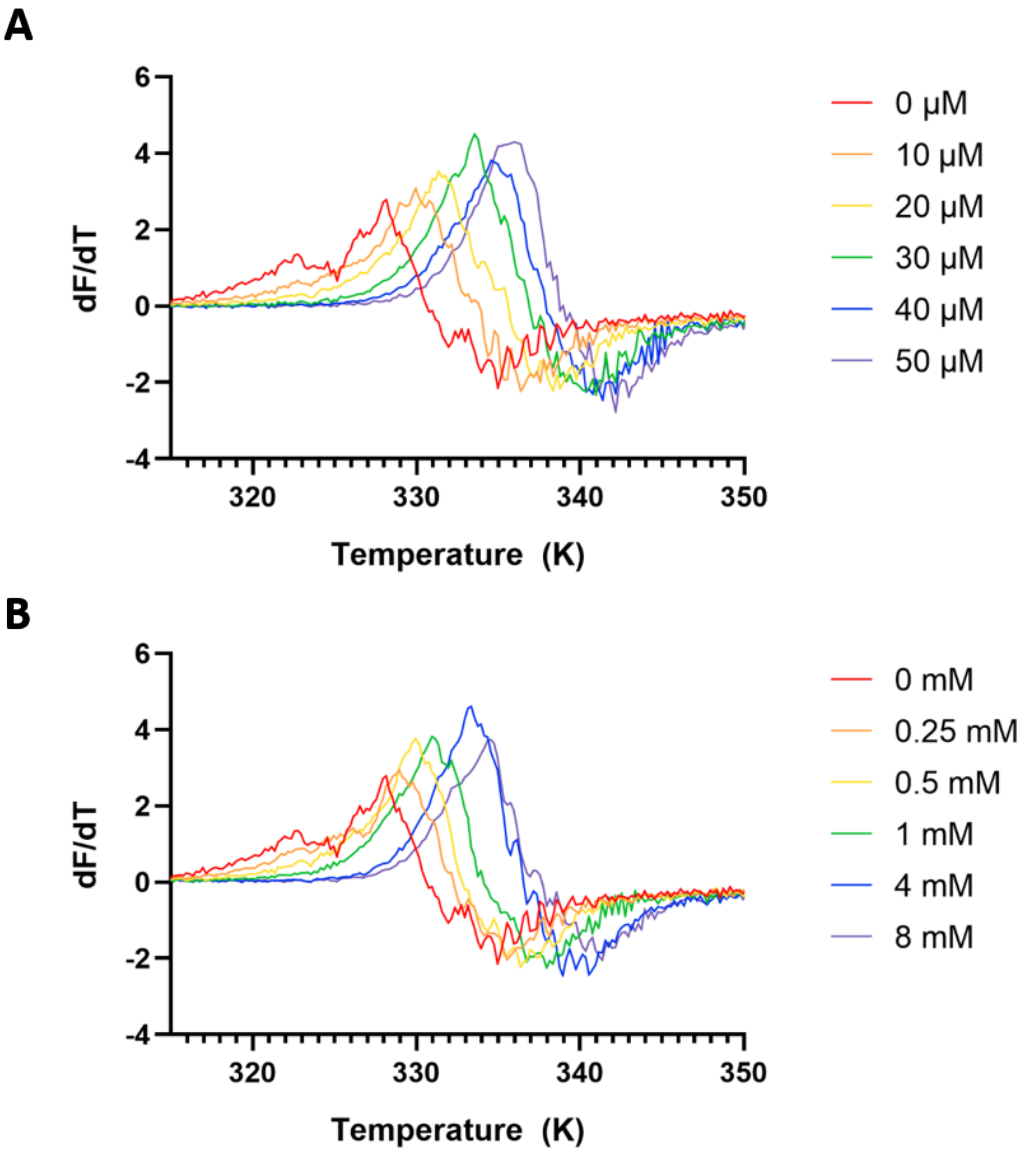

Figure 4.3. First derivative of SYPRO orange thermal shift assay of *mtu*ATP-PRT$_{WT}$ with increasing concentrations of L-Histidine (A), and TIH (B). Both allosteric ligands cause an increasing trend in thermal stability, as shown by increased unfolding temperatures. This increasing trend was saturated by TIH, and nearly saturated by L-His at these concentration ranges. Curves are an average of three replicates.

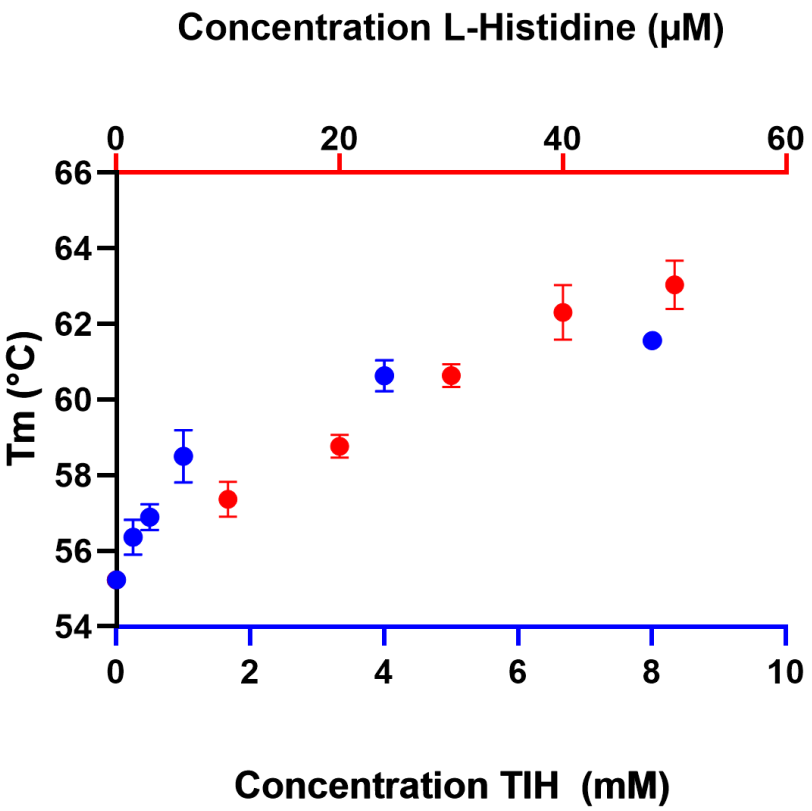

Figure 4.4. Melting temperature (Tm) of *mtu*ATP-PRT$_{WT}$ with increasing concentrations of ligand. Top X axis (Red, L-His (µM)). Bottom X axis (Blue, TIH (mM)). Both L-His and TIH increase the stability of the enzyme with L-His stabilising *mtu*ATP-PRT to a slightly greater extent than TIH. Data is an average of three replicates, error bars where visible are the standard deviation of three replicates.

Both allosteric ligands increase the thermal stability of *mtu*ATP-PRT$_{WT}$ in a similar manner (**Figure 4.4**). This is consistent with previous crystal structures (PDB: 5LHU and 5LHT) solved by Pisco *et al.* (2017) which show the enzyme adopt the **T** state upon binding of both L-His and TIH. This increased temperature stability is also consistent with literature that acknowledges the effect of ligand binding altering enzyme flexibility, and this effect on thermal stability (Celej *et al.*, 2003).

4.2.3 Kinetic assay optimisation

Although there has been significant kinetic characterisation of *mtu*ATP-PRT and ATP-PRT from other species, currently none have been performed with temperature as a variable. As a result, these assays needed additional optimisation for consistency, activity, temperature driven pH shifts, and to counteract variation and artefacts across multiple assays.

Initial temperature data for *mtu*ATP-PRT$_{WT}$ were collected in TRIS (pH 8.5 at 24.5 °C) buffer (**Table 3.10**) consistent with the published ATP-PRT assay systems. Due to the large $\Delta(pK_a)/\Delta T$ of TRIS temperature curves collected showed unusual activity upon binding of L-His (**Figure 4.5**). Initial IC_{50} data under these conditions (Appendix **A.3.1** panel **A**) suggested this to be a phenomenon of the temperature conditions, as the L-His activity collected during the T_{opt} was anomalous. This effect was not reduced by concentration of L-His either (**Figure 4.6**), where inverted temperature curves remained.

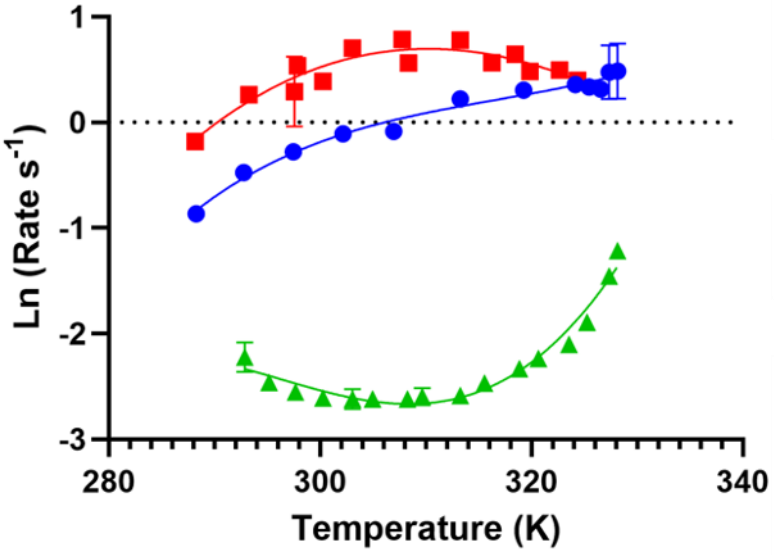

Figure 4.5. Temperature data of *mtu*ATP-PRT$_{WT}$ in a TRIS buffered assay system (**Table 3.10**) with allosteric ligands (L-His 50 µM, TIH 3 mM). APO (blue) and TIH (red) modulated *mtu*ATP-PRT followed typical temperature-log rate kinetics (Hobbs *et al.*, 2013), however L-His (green) bound *mtu*ATP-PRT displays inverted curvature with temperature. This is likely due to a pH shift of 8.6 to 7.4 altering the pronation of L-His and binding abilities. Data is an average of three replicates, error bars where visible are the standard deviation of three replicates.

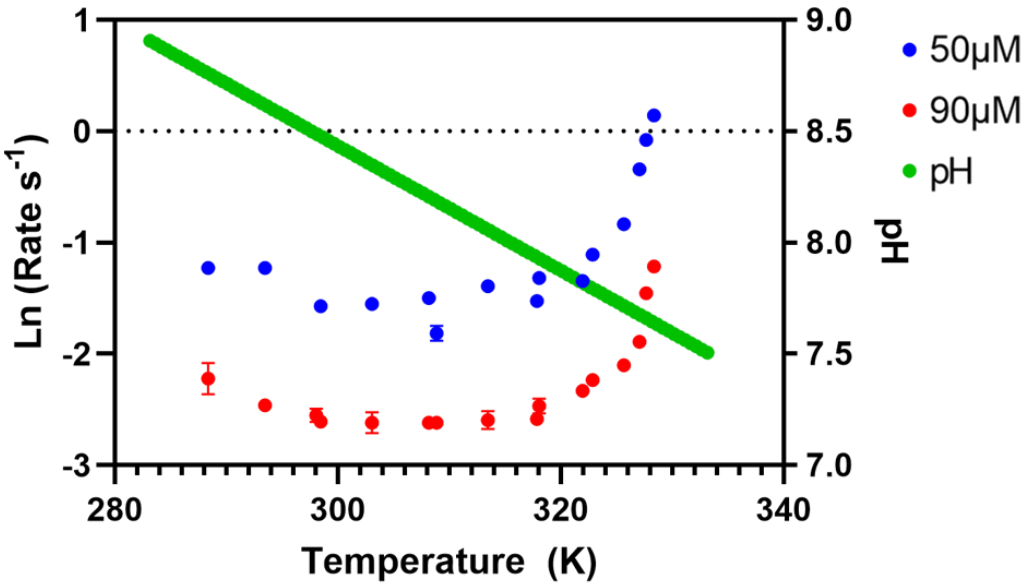

Figure 4.6. T_{opt} assay collected from 293 – 328 K (450 nM *mtu*ATP-PRT$_{WT}$, 3 mM APT, 1.5 mM PRPP, 500 nM *Eco*PPase). The relationship between L-His binding across shifting pH is relative to concentration of inhibitor. L-His begins to loose inhibitory effects at higher temperatures probably as the pH approaches the p*K*a of Histidine (322 K). Data is an average of three replicates, error bars where visible are the standard deviation of three replicates.

Conventionally *mtu*ATP-PRT assays use a TRIS assay buffer (Mittelstädt *et al.*, 2016; Pisco *et al.*, 2017) with a $\Delta(pK_a)/\Delta T$ of -0.028 (**Figure 4.7**). TRIS buffers in an effective range from pH 7.0 – 9.0 and is not known to interact with any metal cofactors. Assaying across a temperature gradient of Δ55 ° (as collected for T_{opt}) exacerbates $\Delta(pK_a)/\Delta T$ of buffers. In TRIS this appears as a pH shift of -1.54, greatly shifting the assay pH over the course of a typical T_{opt} dataset (**Figure 4.7**). This shift in pH was found to alter the binding of L-His across the dataset (**Figure 4.5**, and **Figure 4.6**). Thus, care was taken in selecting a buffering condition with suitable temperature dependence whilst retaining activity. Ideally, data collected over a wide temperature range would use a phosphate buffered assay system with a minimal $\Delta(pK_a)/\Delta T$ of 0.0044 (Hobbs *et al.*, 2013), resulting in a near constant pH across the dataset. However due to the magnesium requirement in ATP-PRT reactions, this buffer is not appropriate, and, as a result we implemented new buffering conditions for ATP-PRT T_{opt} assays.

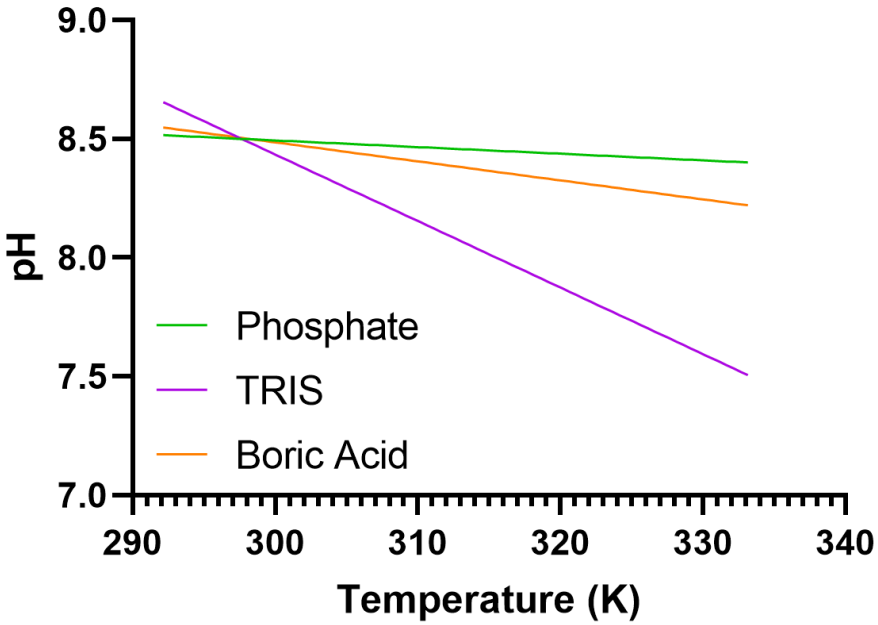

Figure 4.7. $\Delta(pK_a)/\Delta T$ of typical *mtu*ATP-PRT assay buffers (TRIS), phosphate, and boric acid based buffering systems across the temperature range *mtu*ATP-PRT is subject to in temperature assays. $\Delta(pK_a)/\Delta T$ calculated from pH 8.5 at 25 °C.

Boric acid ($\Delta(pK_a)/\Delta T$ of -0.008), phosphate, and TRIS were tested as assay buffers for *mtu*ATP-PRT. Phosphate was unsuitable as a buffering system due to lack of activity, likely driven by inhibition and the phosphate/Mg^{2+} interaction (Ferreira *et al.*, 2015). However, *mtu*ATP-PRT_{WT} remained active in boric acid, and L-His remained inhibitory (Appendix 3, **Figure A.3.3**). A pH curve using Boric acid was performed with *mtu*ATP-PRT_{WT} showing optimal activity beyond pH 9.0 (**Figure 4.8**). pH 9.0 was selected for assays for increased activity allowing better signal during assays whilst staying within the buffering range of boric acid. Following these results boric acid pH 9.0 was used as the buffering agent for downstream *mtu*ATP-PRT assays.

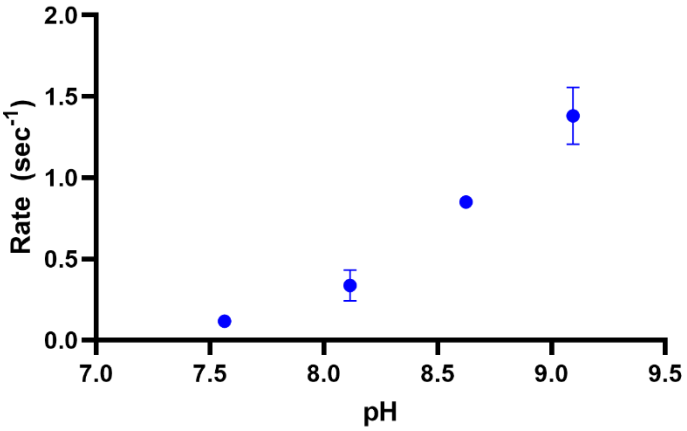

Figure 4.8. Rate of *mtu*ATP-PRT$_{WT}$ against increasing pH in Boric Acid assay buffer (5mM Boric acid, 100mM NaCl, 50mM KCl, 100mM $MgCl_2$, pH 9.0; 450 nM *mtu*ATP-PRT$_{WT}$, 2.5 mM ATP, 680 nM *Eco*PPase, 1.8 mM PRPP). *mtu*ATP-PRT$_{WT}$ was active in this assay system, with an optimal pH of 9.0 or higher. Boric acid does not buffer above pH 9.0, thus this plot was not taken higher in pH. Data is an average of three replicates, error bars where visible are the standard deviation of three replicates.

4.2.4 Binding constant characteristics of *mtu*ATP-PRT and Δ258 mutants

Michaelis–Menten plots (**Figure 4.9**, **Figure 4.9**, and **Figure 4.10**) were collected to determine K_M and k_{cat} values to guide T_{opt} conditions specific to the enzyme. Due to the nature of temperature dependent assay conditions, Michaelis–Menten plots were performed at both 298 and 313 K (25 and 40 °C).

Michaelis–Menten plots of *mtu*ATP-PRT$_{WT}$ (**Figure 4.9**) were collected at two temperatures in the presence and absence of L-His (50 µM). These data fit standard Michaelis–Menten kinetics and show increased catalysis relative to temperature, both in APO and L-His inhibited form. The degree of allosteric inhibition by L-His was exacerbated at higher temperatures with a 40 % decrease in activity at 313 K versus 21 % decrease in activity at 298 K when normalised between modulated and APO k_{cat}.

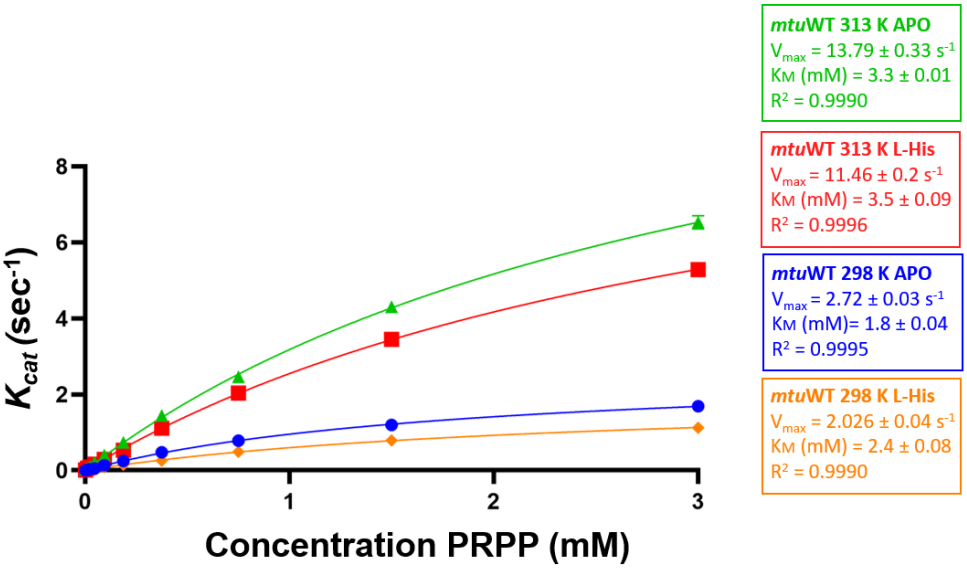

Figure 4.9. Michaelis–Menten plots for PRPP at 298 and 313 K of *mtu*ATP-PRT$_{WT}$ with and without L-His (50 µM). ATP was held at excess (3 mM), *mtu*ATP-PRT$_{WT}$ (450 nM), and *Eco*PPase (500 nM). The addition of L-His reduced enzyme rate at both temperatures however allosteric inhibition did not alter the concentration of substrate needed for saturation. Data is an average of three replicates, error bars where visible are the standard deviation of three replicates.

Michaelis–Menten kinetics for *mtu*ATP-PRT$_{D258N}$ variants at 313 K (**Figure 4.10**) show variation to WT kinetics. Notably *mtu*ATP-PRT$_{D258N}$ showing resistance to L-His inhibition, with inhibited L-His rates aligning similarly to APO. Observing rates alone, we see Δ258A (**Figure 4.10 B)** has both reduced catalysis and reduced K_M, while Δ258N (**Figure 4.10 C**) has increased catalysis compared to WT. This altered rate effect is reflected in each V_{max} value. These data also indicate the reduced inhibitory effect of L-His on Δ258N, with V_{max} values between APO and L-His differing by only 0.15 s^{-1}. Δ258A did not reach saturation under the same conditions as WT and Δ258N (altered K_M), and as a result of this and fitting to Michaelis–Menten kinetics the K_M has been extrapolated and we see large error associated with this. For each variant the K_M is increased slightly with the increase to temperature (**Figure 4.9**).

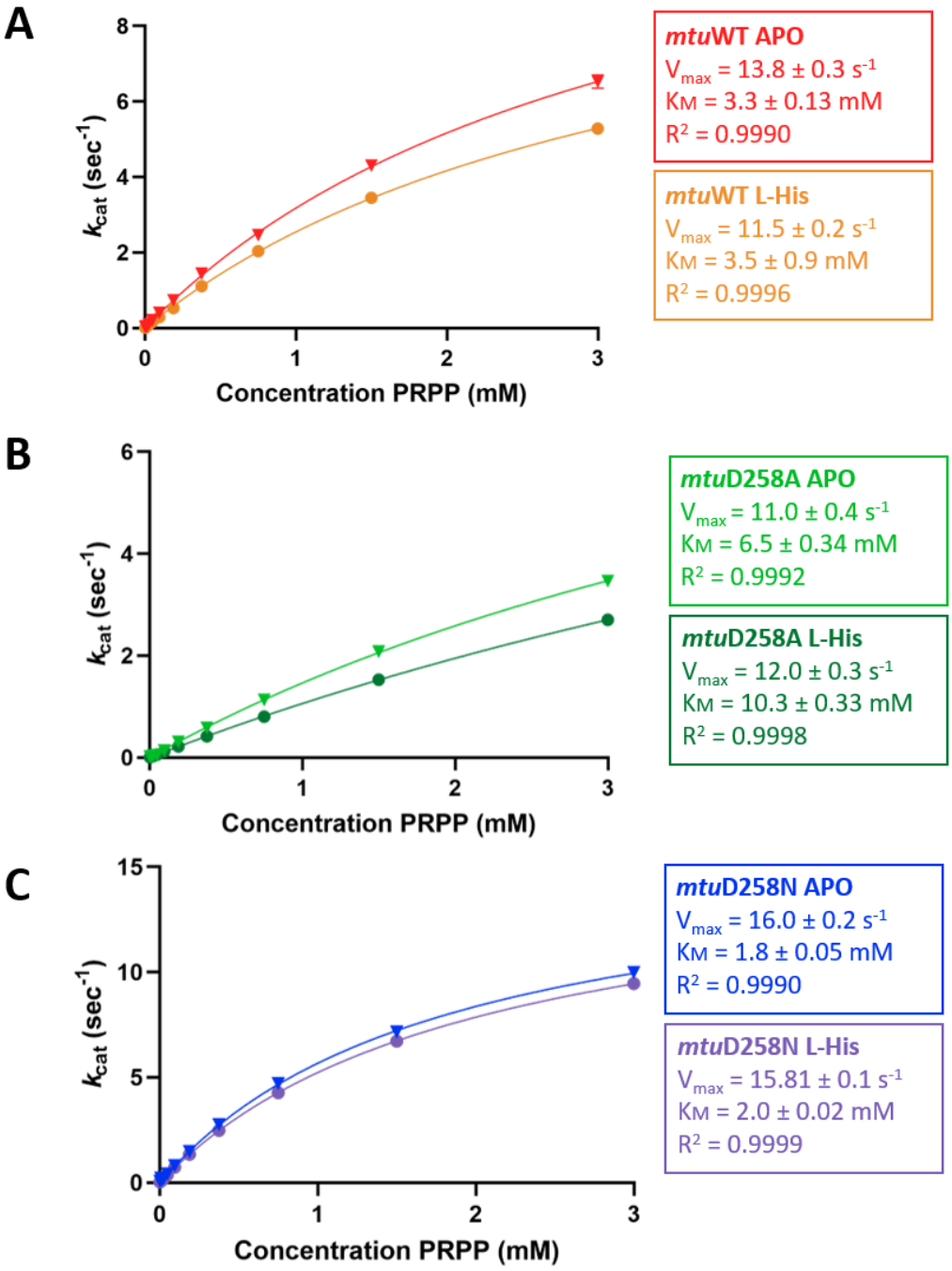

Figure 4.10. Michaelis–Menten plots of *mtu*ATP-PRT$_{WT}$, *mtu*ATP-PRT$_{D258A}$, and *mtu*ATP-PRT$_{D258N}$ at 313 K with ATP saturating at 3 mM. APO WT showed higher catalytic rates compared to ATP-PRT under L-His inhibition, with a slightly decreased K_M for PRPP. *mtu*ATP-PRT$_{D258A}$ APO and L-His inhibited showed significantly increased KM for PRPP compared to other *mtu*ATP-PRT variants and appear unsaturated by PRPP. *mtu*ATP-PRT$_{D258N}$ showed the highest catalytic rates of all *mtu*ATP-PRT variants, with significantly reduced allosteric inhibition by L-Histidine. Each ATP-PRT was at a final concentration of 450 nM, with the addition of *Eco*PPase (500 nM). Data is an average of three replicates, error bars where visible are the standard deviation of three replicates.

K_M values for ATP in **Figure 4.11 B**, collected at 298 K in pH 8.5 TRIS aligned with previously published *mtu*ATP-PRT_{WT} data (Pedreño *et al.*, 2012), however, showed increased V_{max} (1.2 vs 0.31 s^{-1}). Compared to this, in Boric acid, with increased temperature and pH *mtu*ATP-PRT_{WT} displayed greater than 10-fold increased rate, and 34 times increased K_M for ATP. Data for ATP K_M has been extrapolated and this is reflected in the high K_M and error value. This Michaelis-Menten was collected up to 12 mM ATP (Appendix **Figure A.3.4**), however this appeared anomalous, likely due to a lack of Mg2+. Intermediate data points between 6 – 12 mM could not be collected to confirm substrate inhibition due to PRPP supply.

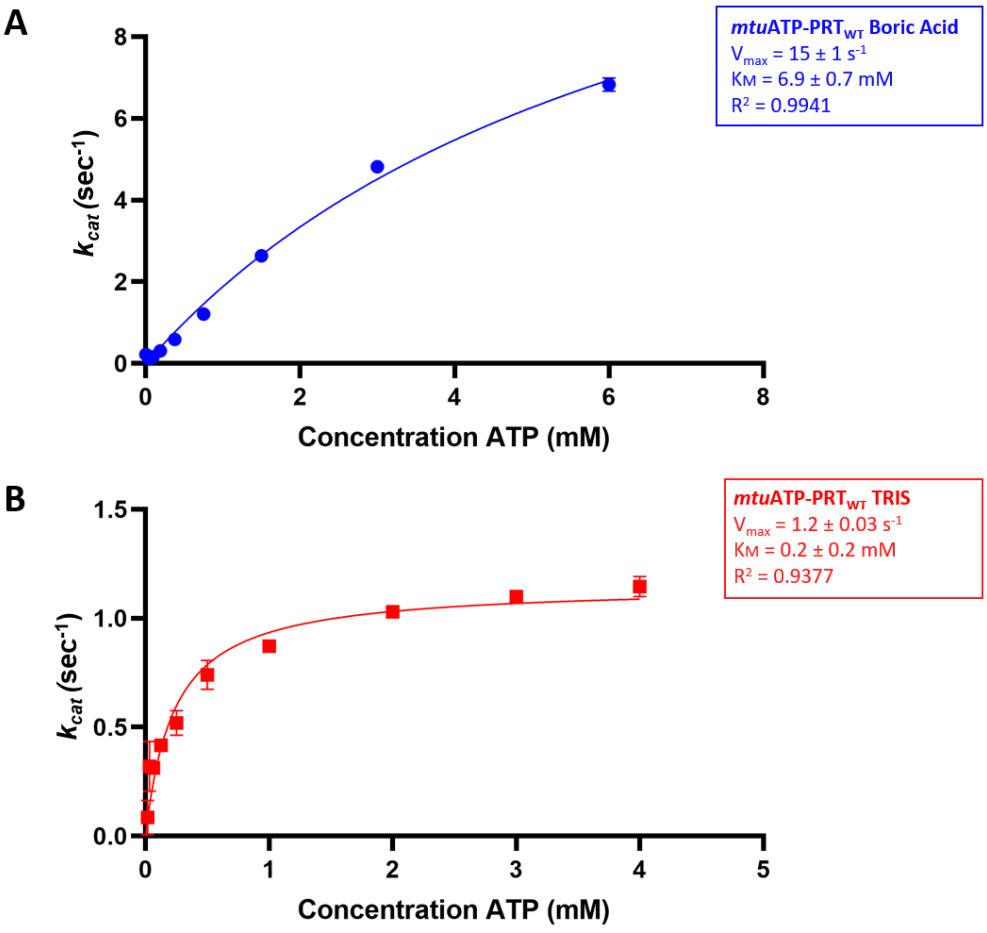

Figure 4.11. Michaelis–Menten plot for ATP. A. Michaelis–Menten collected in final Boric acid buffer with PRPP held at 1.5 mM (*mtu*ATP-PRT_{WT} 450 nM, *Eco*PPase 500 nM). Due to PRPP limitations this plot was only collected for *mtu*ATP-PRT_{WT} at a single temperature of 313 K without the addition of allosteric inhibitors. B. Michaelis–Menten plot of *mtu*ATP-PRT_{WT} collected using TRIS (pH 8.5) assay buffer collected at 298 K. PRPP was held at 1.5 mM. Data for **A** is an average of three replicates, data for **B** is an average of five replicates. Error bars where visible are the standard deviation of respective replicates.

Various Michaelis–Menten plots (**Figure 4.9**, **Figure 4.10**, and **Figure 4.11**) were used to guide substrate concentrations for T_{opt} assays in combination with published literature of ATP-PRT enzymes. The above plots illustrate trends of enzyme rate across temperature, and the effect of allosteric inhibition with the addition of temperature which is later used to investigate temperature effects.

4.2.5 Temperature dependent kinetics of *mtu*ATP-PRT variants

Following the optimisation of *mtu*ATP-PRT variants expression and assay systems, temperature curves were collected via spectroscopy and fit to MMRT models (Section 3.7.3.3). Enzyme reactions were set up in dual syringes; with final concentrations of ATP-PRT (450 nm), *Eco*PPase (500 nM), ATP (3 mM), L-His (25 µM), and TIH (3 mM) in Boric acid assay buffer (pH 9.0 at 298 K). When required L-His and TIH were added to the syringe containing enzyme. All assays were initiated by PRPP (1.5 mM) which was contained in an independent syringe to prevent any reaction mixing before the initiation of experiment. T_{opt} datasets were collected from 288.75 to 333.15 K. These conditions had been used in preliminary T_{opt} data collection and were sufficient in providing enough signal to measure data, whilst being able to show inhibition by L-His. Data were collected for up to one minute and during analysis rate data were manually extracted from the linear portions. Typically, this was between 10 and 20 seconds.

Temperature rate data are presented in **Figure 4.12**. Figures A, B, and C show a good fit to both the temperature independent and temperature dependent MMRT models with traditional enzyme curvature observed across the temperature profile.

Temperature-rate data of *mtu*ATP-PRT_{WT} (**Figure 4.12 A**) shows crossover of APO and L-His reactions at lower and upper temperatures indicating a non-linear temperature relationship. At the lowest collected temperature (288 K) these rate data points coincide, they diverge afterwards and converge at approximately 325 K. TIH modulated *mtu*ATP-PRT_{WT} is faster across most temperatures and does not cross over with any data across the temperature range measured.

Temperature rate data of *mtu*ATP-PRT$_{D258A}$ (**Figure 4.12 B**) shows reduced rates and altered effects of L-His and TIH. TIH still acts as an activator, stimulating increased catalysis across all temperature values, however there is increased curvature in Δ258A causing the TIH curve to meet the L-His and APO curves at 333 K. L-His continues to act as an inhibitor in *mtu*ATP-PRT$_{D258A}$ with altered curvature, showing high levels of inhibition at lowered temperatures (288 K), however this is reduced with increasing temperatures, and at 325 K this inhibitory effect is lost as the temperature curves cross over, and L-His shows near similar or increased rates to the APO curve.

*mtu*ATP-PRT$_{D258N}$ (**Figure 4.12 C**) shows increased catalysis, and lesser allosteric inhibition by L-His. The steepness of all three temperature curves is increased compared to figures **A** and **B**, and the distance between APO and L-His curves is reduced. All three curves track similarly across the entire temperature profile, with the TIH bound curve remaining independent from APO and L-His across the temperature gradient. Similar (**Figure 4.12 A** and **B**) to the merging of APO and L-His curves at higher temperatures, with these two curves meeting approximately 320 K.

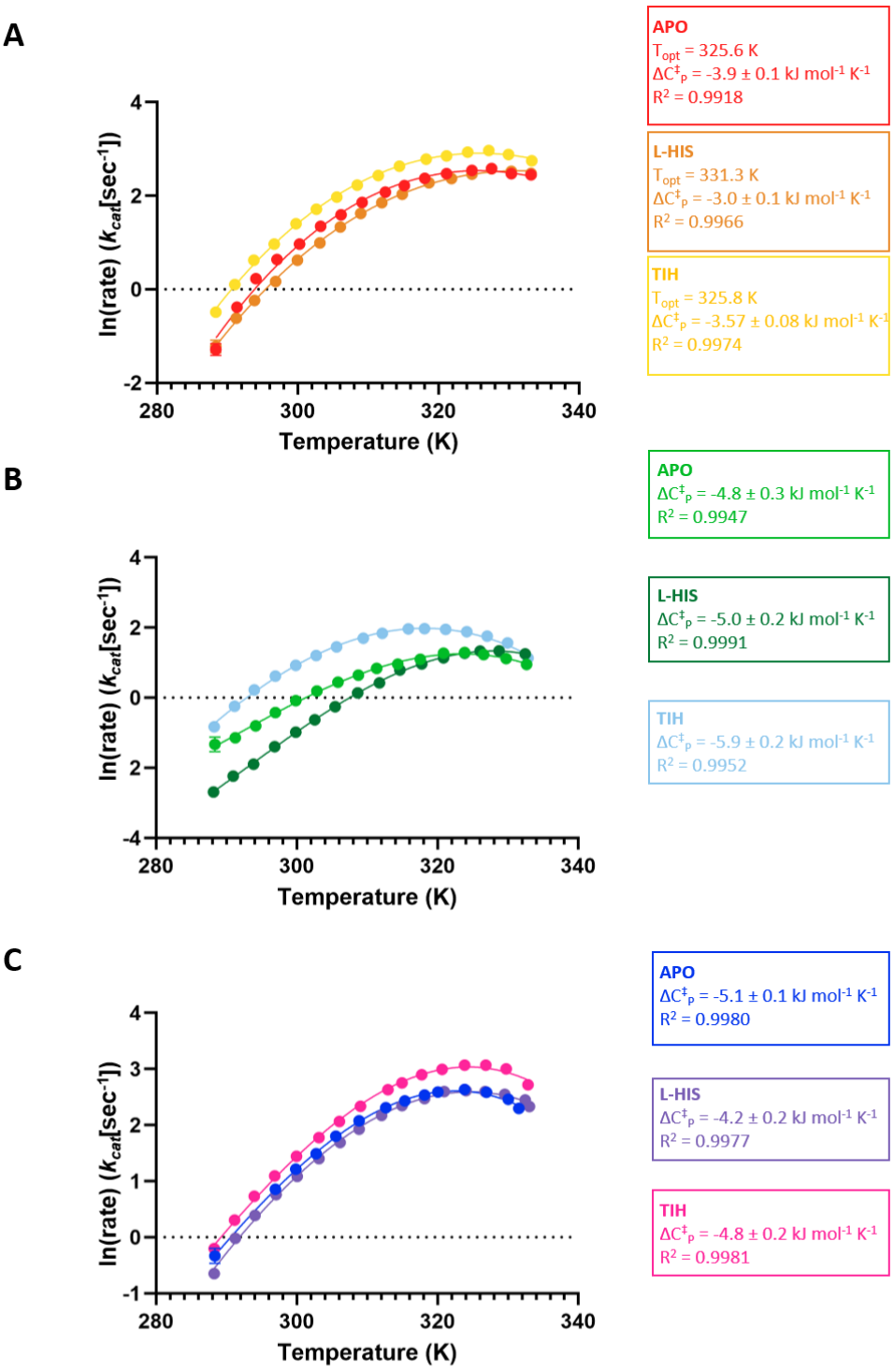

Figure 4.12. Figures A, B, and C show data of *mtu*ATP-PRT variants steady state kinetics collected over a temperature range of 288.15 and 333.15 K with no allosteric modulator, 25 µM L-His, or 3 mM TIH. MMRT models fit with a T_0 of 320 K. Figure A shows APO, L-His and TIH temperature curves of *mtu*ATP-PRT$_{WT.}$ Figure B shows APO, L-His and TIH temperature curves of *mtu*ATP-PRT$_{D258A.}$ Figure C shows APO, L-His and TIH temperature curves of *mtu*ATP-PRT$_{D258N.}$ Data are the average of three replicates, error bars (where visable) are standard deviation of replicates. $\Delta C_p^{\ddagger}$ (at T_0) is given in kJ mol^{-1} K^{-1} for simplicity.

The data illustrated in **Figure 4.12** were fit to the temperature independent (**Figure 4.12**, **A**) and temperature dependent (**Figure 4.12**, **B** and **C**) MMRT models to determine the temperature relationship of the data and extract additional parameters to investigate allosteric regulation in *mtu*ATP-PRT. Each data set were fit to temperature dependent and temperature independent MMRT models using a T_0 of 320 K detailed in **Table 4.1**. 95% confidence intervals and the standard error for each fit parameter were calculated and listed. Akaike information criterion (AICc) (Akaike, 1974) values were used to assess the temperature independent or temperature dependent models of MMRT for further analysis. AICc values for all three WT datasets were similar between MMRT models suggesting the additional parameter of the temperature dependent model did not offer extra explanatory power given the data, and thus further analysis used the temperature independent model. For Δ258A and Δ258N, AICc values suggest introduction of the extra parameter of temperature dependence of $\Delta C_p^{\ddagger}$ adds to the description of the data. Further analysis was undertaken using the temperature dependent model, and respective fit values. These data are shown in **Table 4.2**.

Based on AICc values from (**Table 4.1**), temperature curves for each dataset were generated for appropriate MMRT models. **Figure 4.12 A** shows temperature data from *mtu*ATP-PRT$_{WT}$ fit to the temperature independent MMRT model. Comparing APO and L-His inhibited dataset there is crossover of data at the lower and upper temperature limits. Evident in the reduced curvature of data for L-His inhibition, we see a reduction in $\Delta C_p^{\ddagger}$ for allosteric regulation. The $\Delta C_p^{\ddagger}$ value reports directly on the curvature of the temperature-ln(rate) data. TIH is an activator at all temperatures, in TIH activated *mtu*ATP-PRT$_{WT}$ data sits above APO and L-His data, with similar curvature as indicated by its intermediate $\Delta C_p^{\ddagger}$ value.

Figure 4.12 B shows the temperature curves of Δ258A fit to the temperature dependent MMRT model (with the reference T_0 value of 320 K). Δ258A shows lowered activity compared to WT and Δ258N ATP-PRT variants, yet the change in dynamic properties is significant. Comparing APO and L-His inhibited dataset there is crossover of data at the upper temperature limits, similar to WT. However, there is no cross over at the lower limit, with L-His showing strong inhibition below 310 K. Intriguingly, at 325 K L-His loses inhibitory properties and begins to exhibit activating properties that align with TIH activation at high temperature. TIH

activated Δ258A has a more negative $\Delta C_p^{\ddagger}{}_{T0}$ compared to both the APO and L-His data. This altered $\Delta C_p^{\ddagger}{}_{T0}$ (-5.91 $kJ.mol^{-1}.K^{-1}$) alters the shape of temperature curve bringing the lower and higher ends of the data closer to the APO data. Altering the $\Delta C_p^{\ddagger}$ achieves strong activation in the mid-range temperatures, but this is sacrificed by a reduction in activity at the temperature extremes.

Finally, **Figure 4.12 C** shows the temperature curves of Δ258N fit to the temperature dependent MMRT model (with a fixed T_0 value of 320 K). APO and L-His datasets align very closely across the whole temperature gradient indicating reduced sensitivity to L-His at all temperatures. Thus, this mutation has effectively eliminated L-His inhibition across the temperature range. L-His binding reduces the $\Delta C_p^{\ddagger}$ of Δ258N, reducing the flexibility of the enzyme whilst retaining near WT activity. With TIH activation the temperature data from Δ258N does not stand out from the APO and L-His data. These data sit slightly above APO and L-His, representative of increased catalysis, but demonstrates similar curvature and features no crossing over of catalytic rates within the temperature range collected. Similar to WT, the $\Delta C_p^{\ddagger}{}_{T0}$ of TIH activated is slightly less negative than APO, suggesting lowered flexibility of the enzyme in this state.

Table 4.1. Temperature independent (unshaded) and temperature dependent (shaded) MMRT fits of *mtu*ATP-PRT variants fit at fixed T_0 of 320 K (mean T_0). Given units for; $\Delta C_p^{\ddagger}{}_{T0}$ (J mol $^{-1}$ K $^{-1}$), $\Delta H^{\ddagger}{}_{T0}$ (J mol $^{-1}$), $\Delta S^{\ddagger}{}_{T0}$ (J mol^{-1} K^{-1}), A (kJ mol^{-1} K^{-2}). Standard error is displayed for $\Delta C_p^{\ddagger}$, $\Delta H^{\ddagger}$, $\Delta S^{\ddagger}$, and A.

Mutant	Ligand	$\Delta C^{\ddagger}{}_{pT0}$		$\Delta H^{\ddagger}{}_{T0}$		$\Delta S^{\ddagger}{}_{T0}$		A		R^2	AICc	Topt (K)	Tinf (K)
WT	APO	-3900.0	± 100.0	19000	± 2000	-165	± 6			0.9918	-204	325.6	309.9
		-2600.0	± 300.0	18000	± 2000	-169	± 5	43	± 11	0.9938	-215.1		
	HIS	-3000.0	± 100.0	32000	± 1000	-125	± 4			0.9966	-246.3	331.6	313.3
		-3400.0	± 300.0	33000	± 1000	-124	± 4	-12	± 8	0.9968	-246.1		
	TIH	-3570.0	± 80.0	18000	± 1000	-166	± 3			0.9974	-270.8	325.8	309.4
		-3500.0	± 200.0	18000	± 1000	-166	± 3	2	± 7	0.9974	-268.4		
D258A	APO	-2900.0	± 100.0	15000	± 2000	-189	± 6			0.9878	-211.3	326.1	307.8
		-4800.0	± 300.0	15000	± 1000	-189	± 4	-63	± 8	0.9947	-248.1		
	HIS	-2800.0	± 100.0	44000	± 2000	-100	± 5			0.9949	-211.3	336.7	317.5
		-5000.0	± 200.0	44000	± 1000	-99	± 2	-73	± 5	0.9991	-290.4		
	TIH	-4800.0	± 100.0	-14000	± 1000	-275	± 4			0.9917	-239.3	317.6	303.9
		-5900.0	± 200.0	-14000	± 1000	-272	± 3	-39	± 7	0.9952	-263.1		
D258N	APO	-3600.0	± 100.0	1000	± 1000	-194	± 4			0.9925	-197	321.0	304.9
		-5100.0	± 200.0	11000	± 1000	-191	± 2	-49	± 5	0.998	-244.9		
	HIS	-3420.0	± 80.0	17000	± 1000	-173	± 3			0.9963	-276.8	325.8	309.0
		-4200.0	± 200.0	18000	± 1000	-169	± 2	-30	± 6	0.9977	-298.1		
	TIH	-3300.0	± 100.0	19000	± 1000	-162	± 4			0.9947	-239	326.6	309.4
		-4800.0	± 200.0	20000	± 1000	-159	± 3	-50	± 6	0.9981	-285		

Table 4.2. Abridged Temperature independent (unshaded) and temperature dependent (shaded) fit parameters selected for each dataset based on AICc values in **Table 4.1**. Lower inflection point values (Tinf) displayed. Given units for $\Delta C_p^{\ddagger}$ (J mol $^{-1}$ K $^{-1}$), $\Delta H^{\ddagger}$ (J mol $^{-1}$), $\Delta S^{\ddagger}$ (J mol^{-1} K^{-1}), A (kJ mol^{-1} K^{-2}), lower inflection point (Tinf) (K). Each fit is at a fixed T_0 of 320 K.

Mutant	Ligand	$\Delta C^{\ddagger}_{p\ T0}$	$\Delta H^{\ddagger}_{T0}$	$\Delta S^{\ddagger}_{T0}$	A	R^2	Topt (K)	Tinf (K)
WT	APO	-3900.0	19000	-165		0.9918	325.6	309.9
	HIS	-3000.0	32000	-125		0.9966	331.3	313.3
	TIH	-3570.0	18000	-166		0.9974	325.8	309.4
D258A	APO	-4800.0	15000	-189	-63	0.9947		
	HIS	-5000.0	44000	-99	-73	0.9991		
	TIH	-5900.0	-14000	-272	-39	0.9952		
D258N	APO	-5100.0	11000	-191	-49	0.998		
	HIS	-4200.0	18000	-169	-30	0.9977		
	TIH	-4800.0	20000	-159	-50	0.9981		

Examining extracted MMRT fit values (**Table 4.2**) there is significant variation in temperature of maximum rate values between variants and allosteric regulated states from 318.2 to 327.6 K. $\Delta C_p^{\ddagger}$ also varies significantly (by up to 2.8 kJ mol^{-1} K^{-1}), with the highest $\Delta C_p^{\ddagger}$ observed in WT (L-His), and lowest $\Delta C_p^{\ddagger}$ in Δ258A (TIH) bound.

$\Delta H^{\ddagger}_{T0}$ of Δ258A (TIH) is significantly lower than the $\Delta H^{\ddagger}_{T0}$ of all other datasets, which is the result of setting the reference temperature, T_0, at 320 K. Only in the case of Δ258A (TIH) is this reference temperature above T_{opt} leading to a negative value of $\Delta H^{\ddagger}_{T0}$ here ($\Delta H^{\ddagger}$ passes through zero close to T_{opt}). For ease of comparison, 320 has been used as T_0 for all data sets. For example, when this dataset is fit using a T_0 of 314 (4 K below the temperature of maximal rate) the temperature dependent $\Delta H^{\ddagger}_{T0}$ value becomes positive at 19726 J mol^{-1} (Appendix 3, **Table A.3.1**).

In this dataset we observe an entropy-enthalpy trade-off. There is a near-linear relationship between $\Delta H^{\ddagger}_{T0}$ and $\Delta S^{\ddagger}_{T0}$ as illustrated in the x-y plane in **Figure 4.13**. The $\Delta H^{\ddagger}_{T0}$ and $\Delta S^{\ddagger}_{T0}$ values of Δ258A with L-His and TIH bound sit at the extreme edges of **Figure 4.13** and this is consistent with these two displaying the extremes of inhibition and activiation, respectively. WT (L-His) and Δ258A (TIH) sit at the extremes of the $\Delta C_p^{\ddagger}{}_{T0}$ values which is also consistent with changes in rigidity as indicated by $\Delta S^{\ddagger}$ (i.e. from the $\Delta H^{\ddagger}$-$\Delta S^{\ddagger}$ trade-off). There is a good correlation between $\Delta C_p^{\ddagger}{}_{T0}$ and $\Delta S^{\ddagger}$ in **Figure 4.13** with the exception of *mtu*ATP-PRT_{D258A} L-His which is an outlier.

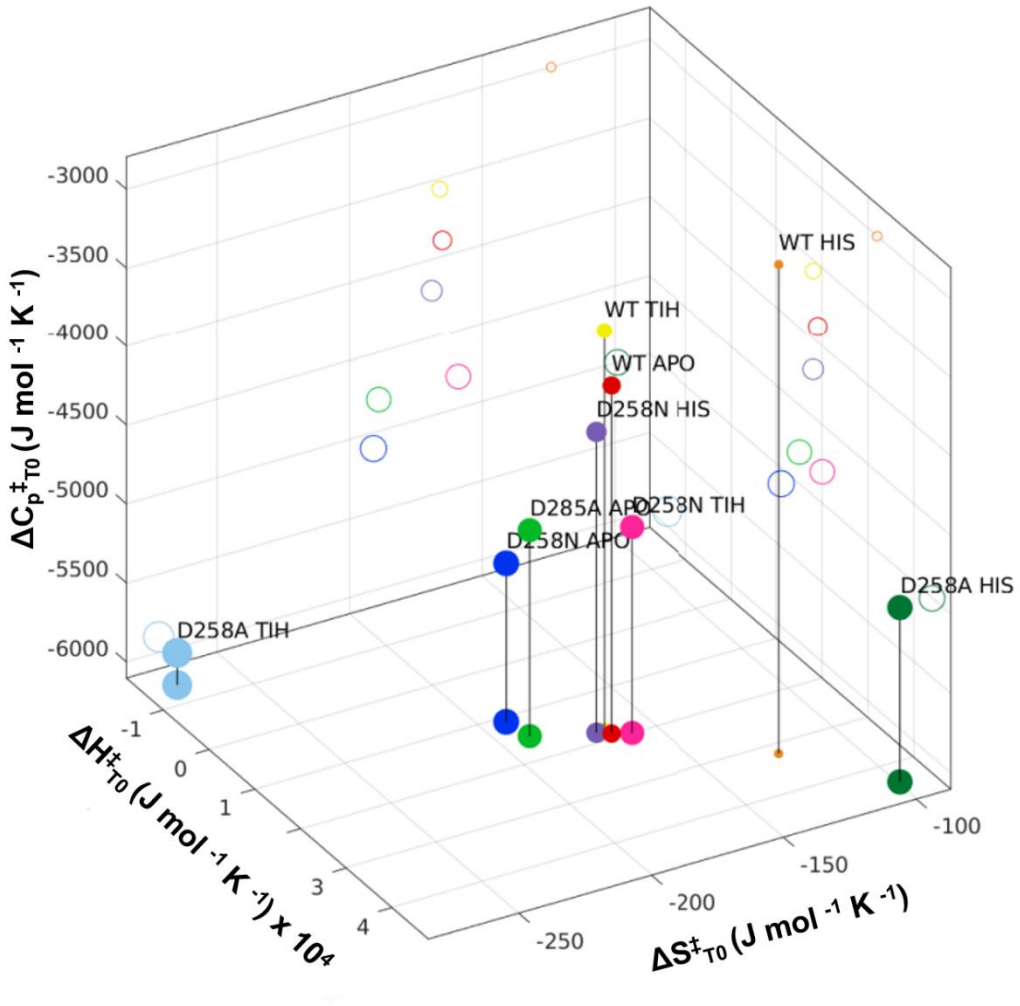

Figure 4.13. 3D graph illustrating the relationship of $\Delta H^{\ddagger}$, $\Delta S^{\ddagger}$, and $\Delta C_p^{\ddagger}$ at T_0 (320 K) extracted from **Table 4.2**. The strong linear trend shows a clear entropy enthalpy trade off, underlying the effects governing the temperature dependence of ATP-PRT mutants and allosteric effector temperature dependence. A result of increasing $\Delta H^{\ddagger}$ is a less negative $\Delta S^{\ddagger}$. As the $\Delta S^{\ddagger}$ becomes more negative, the $\Delta C_p^{\ddagger}$ value also becomes more negative and forms a linear trend between $\Delta S^{\ddagger}$ and $\Delta C_p^{\ddagger}$.

The A factor (**Table 4.2**) represents the temperature dependent slope of the $\Delta C_p^{\ddagger}$ and is derived only from the temperature dependent equation. That is, for the temperature independent fits of WT, A is zero, and the $\Delta C^{\ddagger}_p$ is the same at all temperatures. This slope formulates the temperature dependent relationship of $\Delta C^{\ddagger}_p$, and further alters the temperature relationship of $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$. This can best be observed in **Figure 4.14** (temperature independent) and **Figure 4.15** (temperature dependent). In **Figure 4.14** $\Delta C^{\ddagger}_p$ remains the same over temperature, thus $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ are linearly dependent upon temperature. In **Figure 4.15** $\Delta C^{\ddagger}_p$ is temperature dependent, and as a result changes the temperature dependence of $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ illustrated by curved $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ values across the temperature range.

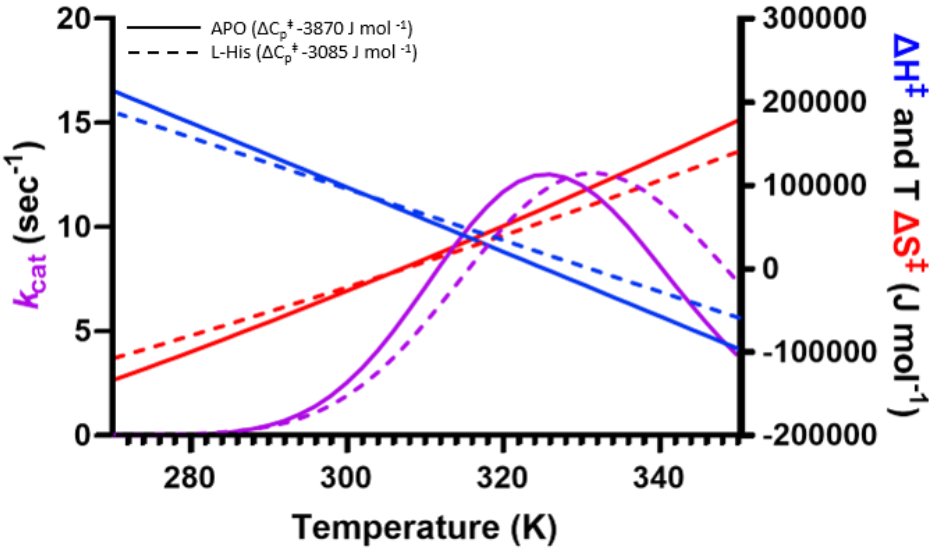

Figure 4.14. Temperature dependence of fit parameters for *mtu*ATP-PRT$_{WT}$ from **Table 4.2**. Rate (Purple), $\Delta H^{\ddagger}$ (Blue), $\Delta S^{\ddagger}$ (Red). For APO (solid line), and L-His (dashed line).

This temperature independent model shows linear $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ relationships across temperature as the $\Delta C_p^{\ddagger}$ value is constant across the temperature range. **Figure 4.14** (solid lines) APO *mtu*ATP-PRT$_{WT}$ illustrates a standard rate curve with a temperature optimum around 325 K. $\Delta H^{\ddagger}$ crosses through zero close to T_{opt} (328 K). **Figure 4.14** (dashed line) shows similar rate curvature to the APO state, however, the addition of L-His shifts the T_{opt} of this curve, up to 328 K. This is a direct result of a less negative $\Delta C_p^{\ddagger}$ value for L-His-inhibited enzyme (*i.e.* reduced curvature in the temperature-rate plot) and reduced slopes of the $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ lines.

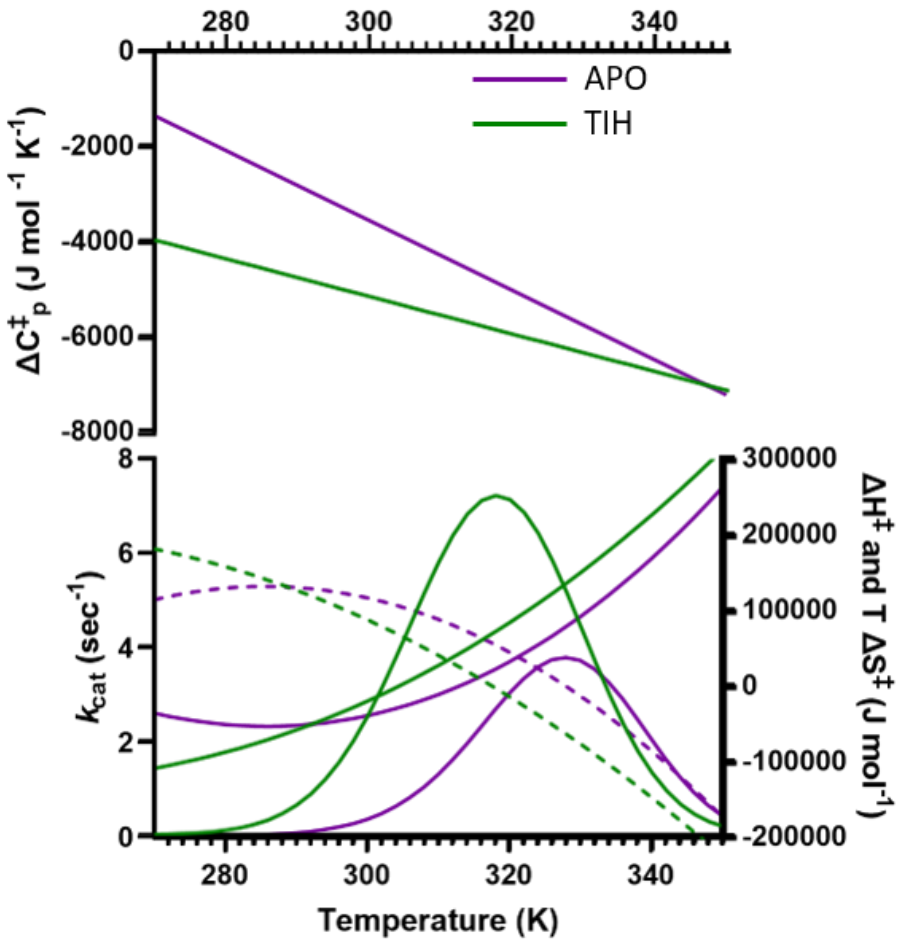

Figure 4.15. Extrapolated temperature dependence of fit parameters for *mtu*ATP-PRT$_{D258A}$ APO (purple) and TIH (green) from **Table 4.2**. $\Delta H^{\ddagger}$ is decreasing and differentiated by dashed lines. $\Delta Cp^{\ddagger}$ is linearly dependent on temperature, and $\Delta S^{\ddagger}$ and $\Delta H^{\ddagger}$ are curved under the temperature-dependent MMRT model (compared to temperature-independent **Figure 4.14**). The differing slopes of $\Delta C_p^{\ddagger}$ between APO and TIH bound states reflects the subsequent curvature in $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$.

This temperature dependent model (**Figure 4.15**) shows non-linear $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ across temperature as the slope of $\Delta C_p^{\ddagger}$ (A) produces a curved temperature dependent relationship compared to the linear temperature dependence of $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ seen in (**Figure 4.14**). These $\Delta S^{\ddagger}$ and $\Delta H^{\ddagger}$ slopes are seen for each temperature dependent model fit of *mtu*ATP-PRT$_{D258A}$ and *mtu*ATP-PRT$_{D258N}$ fits. This temperature dependent model shows non-linear $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ across temperature as the slope of $\Delta Cp^{\ddagger}$ (A) produces a temperature dependent relationship. *mtu*ATP-PRT$_{D258A}$ (APO, and TIH modulated) show similar graphical topology when graphed, driven by their similar $\Delta C_p^{\ddagger}$ and A values (**Table 4.2**). For

*mtu*ATP-PRT$_{D258N}$ (not graphed) a reduced $\Delta H^{\ddagger}$ and $\Delta S\ddagger$ curvature across temperature is driven by the lowered A value of 29.6. The closer the A value gets to zero, the closer the $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ gets to a temperature independent, and thus curvature of $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ is reduced.

Figure 4.12 A B and **C** show points where APO and L-His temperature curves overlap, particularly in lower and upper temperature limits. This is most obvious in temperature curves of Δ258 (**Figure 4.12 B**). The relative increase and decrease in rates (compared to APO rate) in response to TIH and L-His has been graphed in **Figure 4.16 A B** and **C** to illustrate the effect in which each allosteric modulator alters activity compared to APO.

In *mtu*ATP-PRT$_{WT}$ (**Figure 4.16 A**) the relative increase and decrease of rate tracks across the course of the temperature range, merging around 350 K. Like **Figure 4.12 A**, APO and TIH rates do not cross over in the temperature range of the experiment (288 – 333 K). In contrast, in *mtu*ATP-PRT$_{D258A}$ (**Figure 4.16 B**) these relative rates merge around 330 K, and their allosteric effect becomes inverted. Finally *mtu*ATP-PRT$_{D258N}$ (**Figure 4.16 C**) again shows the crossover of rates near 340 K. In mtuATP-PRT$_{D258N}$ T_{opt} data (**Figure 4.12 C**) this cross over is not observed. In this case, T_{opt} data was not collected up to 340 K, and based on the increased curvature observed in *mtu*ATP-PRT$_{D258N}$ data sets we can reasonably expect if the data collection was continued toward 340 K, we would observe a steep decrease to the rate of TIH activated data, which would cross over with L-His bound data thus resulting in an inversion of allosteric effect, similar to that seen in *mtu*ATP-PRT$_{D258A}$ (**Figure 4.16 B**)

Although the relative activity between APO and allosterically regulated rates are an exaggerated version of activity outside of the temperature ranges collected experimentally, they illustrate the trajectory of these allosteric effects over temperature. They also serve to illustrate that the loss of allosteric effect at upper and lower temperature limits is not a function of lack of binding, but an inversion or reduction to allosteric effect. This is to be expected from the thermodynamic considerations illustrated in **Figure 4.13** and the temperature dependence of $\Delta C_p^{\ddagger}$ $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$.

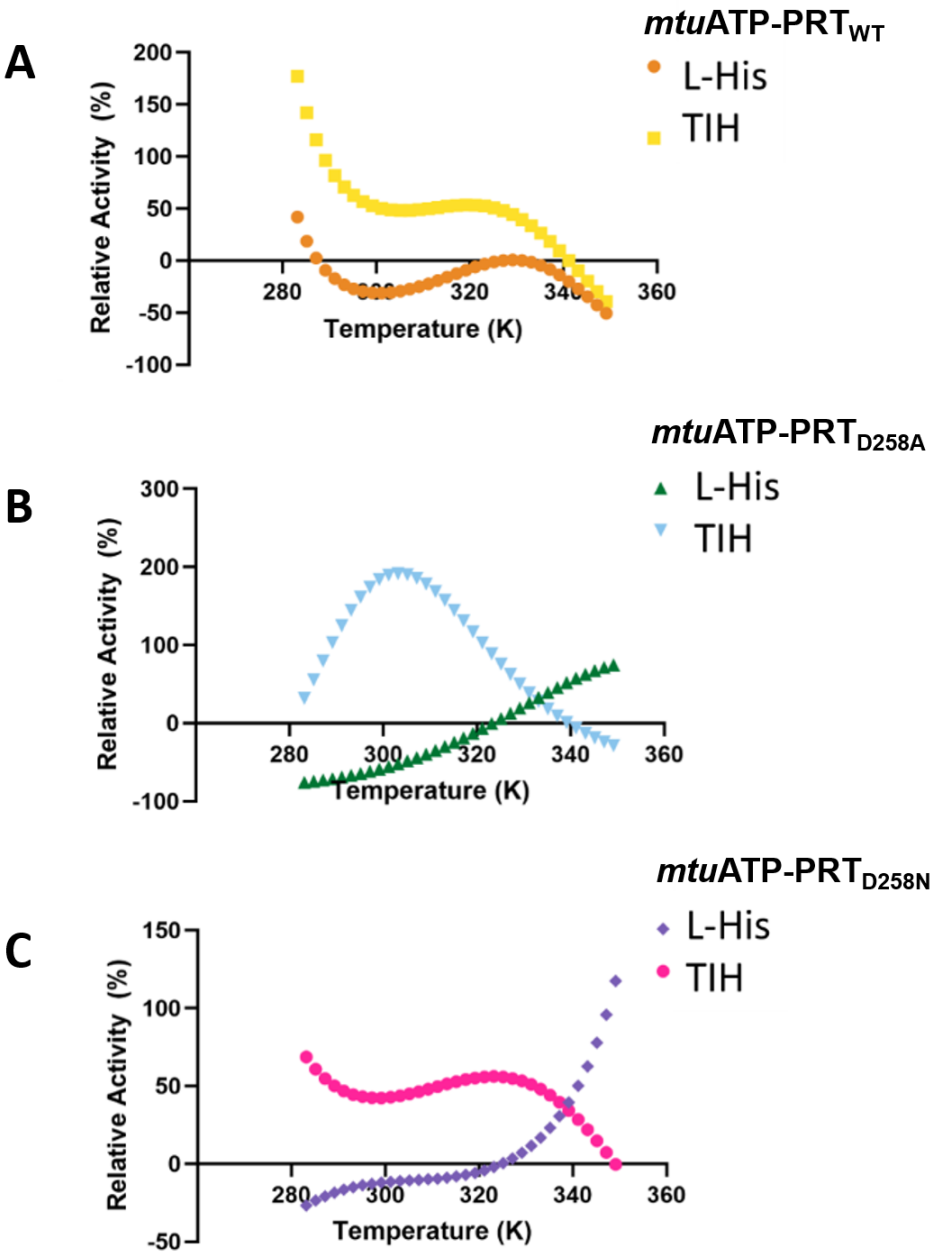

Figure 4.16. Extrapolated relative rates of *mtu*ATP-PRT variants with inhibition by L-His or activation by TIH. **A**. The relative activity (compared to APO rate) changes across temperature following allosteric modulation in *mtu*ATP-PRT$_{WT}$. **B.** The relative activity changes across temperature following allosteric modulation in *mtu*ATP-PRT$_{D258A}$. **C**. The relative activity changes across temperature following allosteric modulation in *mtu*ATP-PRT$_{D258N}$.

4.3 Discussion

The expression and purification of the recombinant protein *mtu*ATP-PRT variants was optimised to produce consistent purification yields of stable and active enzyme. These enzymes were bulk produced and frozen for future use in thermal stability

assays, isothermal titration calorimetry, and steady state kinetics. The ability to generate reproducible yields of enzyme for temperature analysis (both steady state and thermal assay) has allowed us to examine the temperature dependent curvature of enzymatic reactions. Further, this curvature is not due to changes in substrate binding or denaturation of the protein. This section will discuss results of *mtu*ATP-PRT variant kinetics, and MMRT parameters, and will be further expanded in Chapter 7 in the context of crystal structures and molecular dynamics simulations.

4.3.1 Thermal stability of *mtu*ATP-PRT$_{WT}$ indicates tensed conformational states with both allosteric modulators

Differential scanning fluorimetry of *mtu*ATP-PRT$_{WT}$ with both allosteric activator and inhibitor allowed us to assess the thermal stability of *mtu*ATP-PRT in each allosteric conformational state. These results showed both L-His and TIH increased the thermal stability of *mtu*ATP-PRT$_{WT}$, as shown by increasing melt temperatures (**Figure 4.3**). This is a general trend related to binding events due to the increased number of non-covalent bonds in the ligand bound conformation (Celej *et al.*, 2003).

With increasing concentration of the ligands, both saw the thermal stability of the enzyme increase. This was slightly more pronounced with L-His inhibition (336 K *versus* 333 K). Increased thermal stability of *mtu*ATP-PRT in these allosteric bound states likely relates to the tensed conformation the structure adopts, evident in PDB: 5LHU and 5LHT structures produced by (Pisco *et al.*, 2017). With increased numbers of non-covalent bonds from ligand binding, along with increased bonds to elicit a tensed protein state, it is expected protein stability and T_m increases. Both ligands stimulate an increased T_m, accompanied by increased $\Delta C_p^{\ddagger}$ for the WT enzyme (**Table 4.2**), indicating reduced flexibility to the enzyme. Similarly, this effect is more pronounced for L-His. Differential scanning fluorimetry results suggest allosteric ligand binding incites rigidification of the starting state of the enzyme.

It is curious that both ligand bound states elicit very similar increases to thermal stability, accompanied by similar trends in $\Delta C_p^{\ddagger}$ yet one increases catalytic rate whilst the other inhibits the catalytic rate. In reference to $\Delta C_p^{\ddagger}$ (-3900 and -3570 J

mol $^{-1}$ K $^{-1}$ for L-His and TIH respectively) this denotes the $\Delta C^{\ddagger}_{p}$ of the reaction, not the enzyme. Thus it does not necessarily mean that the $\Delta C_p^{\ddagger}$ solely pertains to the ground state, as the transition state flexibility (C_p) may be where the changes driving differential allosteric modulation are occurring.

4.3.2 Assay buffering conditions

Although there is a suite of high-quality data collected for ATP-PRTs using a TRIS buffered assay system, this was unsuitable for temperature-based assays and required optimisation. Care was taken in optimisation to minimise changes to existing assay conditions reported by (Mittelstädt *et al.*, 2016; Pisco *et al.*, 2017), however, many Goods buffers around the pH 8.5 were noted to have either large ΔpKa/ΔT or had innate metal affinities for Mg^{2+} or KCl, necessary for ATP-PRT catalysis. Phosphate buffer has been used with great success in the collection of T_{opt} data for other enzyme systems due to its minimal ΔpKa/ΔT. However, phosphate could not be used due to interactions with $MgCl_2$ critical to ATP-PRT reaction and potential general inhibition of the enzyme *via* phosphate binding. This was tested and no activity was seen with phosphate buffered assay buffers at multiple temperatures.

pH was altered to 9.0 (*c.f.* 8.5) due to the pH range of Boric acid, allowing a single buffer preparation to be used across a temperature assay collection instead of multiple buffers with pH adjusted to 8.5 for each temperature. It was critical to ensure the buffer would stay within its pH buffering range over the course of the temperature range, particularly at the edge temperatures, allowing the capacity for assays to be continued into higher temperature ranges if curvature had not been observed. Additionally T_{opt} assays are ideally collected at the pH optima of the enzyme as rates are less pH dependent and thus pH changes effect rates less. The high pH optima of *mtu*ATP-PRT made this impossible to collect at pH optima whilst staying inside the pH range of appropriate buffers, and thus an intermediate value was chosen. Consequently, the increased pH of the assay buffer moving to reaction closer to the optimal pH of *mtu*ATP-PRT offered the advantage of higher catalytic turnover of the reaction, and thus higher quality data that could be observed, even when activity was lowered by temperature and allosteric inhibition.

We acknowledge that the change to buffering system, and consequently pH adds confounding factors to these assays, and moves the assay system away from both published data, and physiological conditions. TRIS could have been used with individual buffer series; however this would likely introduce significant errors between temperatures.

4.3.3 Kinetic characteristics of mtuATP-PRT

Typical Michaelis–Menten kinetics were observed for *mtu*ATP-PRT$_{WT}$ (**Figure 4.9**). Across the temperature range we observed increased rates, and a modest increases in K_M for PRPP. As a result of increasing K_M over temperature, for data collected in subsequent temperature assays we could expect a decrease in substrate saturation (Thomas & Scopes, 1998). Δ258A (**Figure 4.10** B) had a decreased V_{max} and an increased K_M compared to WT and was unable to reach saturation using the same conditions of other *mtu*ATP-PRT variants. Due to lack of saturation V_{max} and K_M values were extrapolated upon data fitting from collected data and appear very high compared to WT. Δ258N displayed increased V_{max} compared to WT. Along with reduced sensitivity to L-His (**Figure 4.10** C) Δ258N had a decreased K_M for PRPP, thus a higher affinity for PRPP.

Although the Michaelis-Menten curves are well behaved, the calculated K_M values for PRPP are recognised to be unusually high when compared to previously published values (Pedreño *et al.*, 2012). In retrospect, this is most likely due to PRPP degradation in transit. The original strategy to counteract this was to solubilise, aliquot and freeze PRPP from each shipment to ensure consistency across all of the assays. It was not possible to accurately assess PRPP concentrations using *mtu*ATP-PRT and running the reaction to completion. Towards the end of this research we obtained *Lactococcus lactis* ATP-PRT (*lla*ATP-PRT, short form) from our collaborators to assess PRPP concentrations from frozen stocks (by running the reaction to completion) and this suggested PRPP concentrations may be ~5-10 times lower than presented here. This makes the absolute values of K_M reported here likely artificially high. However, it does not affect the temperature dependence curves (**Figure 4.12**) which were collected at a single, consistent concentration of PRPP throughout. Subsequent collection of Michaelis–Menten curves was disrupted due to PRPP supply issues.

4.3.4 Temperature characteristics of *mtu*ATP-PRT

*mtu*ATP-PRT$_{WT}$ fits the temperature independent MMRT model, indicated by fit data, AICc values, and the more complex shape of the temperature rate relationship. Through fitting this model we see changes to the $\Delta C_p^{\ddagger}$ with allosteric inhibition by L-His, suggesting allosteric control may be achieved *via* dynamic changes as indicated by changes in $\Delta C_p^{\ddagger}$. T_{opt} with L-His inhibition also increases with the addition of L-His, suggesting the binding of L-His stabilises the enzyme complex, thus altering the temperature dependence and optimal working temperature of *mtu*ATP-PRT$_{WT}$ through $\Delta C_p^{\ddagger}$ alterations. This notion aligns with data published by Hobbs *et al.* (2013).

Between APO and allosterically activated states there are only slight changes to T_{opt} and $\Delta C_p^{\ddagger}$. In contrast, differences $\Delta S^{\ddagger}_{T0}$ and $\Delta H^{\ddagger}_{T0}$ coordinate with alterations to rate in data extracted from **Figure 4.12** (**Table 4.1** and **Table 4.2**) between variant and allosteric states. These changes represent the enthalpy-entropy trade-off that alters the free energy landscape, and subsequent rate temperature relationship (Arcus & Mulholland, 2020). TIH activated reactions appear to behave like psychrophilic enzymes, with decreasing $\Delta H^{\ddagger}$ and increasing $\Delta S^{\ddagger}$ at T_0. Similarly, the L-His inhibited reactions are behaving similar to thermophilic enzymes, with larger $\Delta H^{\ddagger}$ coupled with less negative $\Delta S^{\ddagger}$.

L-His reduced enzyme rates by some degree in all variants for the majority of the temperature range. All L-His assays showed an increase in $\Delta H^{\ddagger}$ and a decrease in $\Delta S^{\ddagger}$, suggesting that the inhibition involves an enthalpy-entropy trade-off. For TIH, both WT and Δ258A showed decreased $\Delta H^{\ddagger}$ and increased $\Delta S^{\ddagger}$, however this was not seen in Δ258N with TIH bound which had a slightly increased $\Delta H^{\ddagger}$ in favour of a decreased $\Delta S^{\ddagger}$. The strong linear trend observed in **Figure 4.13** for all enzymes shows a clear $\Delta H^{\ddagger}$, $\Delta S^{\ddagger}$ trade off underlying the effects governing the temperature dependence of *mtu*ATP-PRT enzymes and allosteric effector temperature dependence. This trade-off also extends to $\Delta C_p^{\ddagger}$ with only one exception – *mtu*ATP-PRT$_{D258A}$ where L-His has an anomalous more negative $\Delta C_p^{\ddagger}$ value than expected. Further discussed in Chapter 5 (Structural characteristics of allosteric regulation in *mtu*ATP-PRT) in reference to the structure and residue interactions

*mtu*ATP-PRT$_{D258A}$ appears to be a special case, with extreme divergence of inhibition and activation at low temperatures that may provide additional clues to understand the enthalpy-entropy trade-off that drives allostery.

An unexpected feature of *mtu*ATP-PRT$_{WT}$ temperature assays is the meeting of APO and L-His data points at the lower temperature range, and the crossover of these at the high temperature range, with L-His becoming an activator at high temperatures. These data points meet at 328 K and due to the increased L-His bound T_{opt} of 313 K increased activity is seen compared to APO at the higher temperatures. If data were collected beyond 333 K, we may see a cross over between TIH bound and L-His bound too, due to the nature of each curve. Extrapolated data shows the relative activity of inhibited and activated enzyme cross over at temperature extremes. Curvature driven by $\Delta C_p^{\ddagger}$ drives the degree in which this occurs.

4.3.5 Mutation to Asp258 alters allosteric response to L-Histidine

By mutating residue 258 from a negatively charged side chain to an uncharged sidechain (asparagine) we have reduced the effect of L-His inhibition in *mtu*ATP-PRT. Whilst inhibition has been reduced, allosteric activation by TIH remains functional. This suggests although L-His and TIH share chemical similarity, and elicit a tensed response in *mtu*ATP-PRT (Pisco *et al.*, 2017), they may communicate the regulatory message *via* alternate pathways. When this same residue was truncated from aspartic acid to alanine (Δ258A) sensitivity to both L-His and TIH was increased at low temperatures, suggesting a crucial role of residue 258 in the communication of allosteric signal transduction. The rates of *mtu*ATP-PRT$_{D258A}$ were also reduced compared to WT and Δ258N. Discussed in Chapter 5, crystal structures of each of these mutants have been collected with L-His and TIH bound.

In *mtu*ATP-PRT$_{D258N}$ L-His insensitivity is accompanied by a decrease in $\Delta C_p^{\ddagger}$ compared to APO enzyme, and an increase of T_{opt}. Values of $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ do not vary hugely between any of the Δ258N states at T_0 (320 K), however the A factor of L-His bound is less negative, thus altering the temperature dependence of $\Delta S^{\ddagger}$, and $\Delta H^{\ddagger}$, so these parameters diverge for these enzyme systems away from T_0.

*mtu*ATP-PRT$_{D258A}$ illustrated significant changes to the temperature landscape in allosteric states. Δ258A appears to increase the sensitivity to L-His particularly at lower temperatures. However, this is accompanied by a shift in the curvature of the enzyme temperature profile, and as a result L-His becomes an activator at temperatures beyond 324 K. With L-His, and TIH bound $\Delta C_p^{\ddagger}$ becomes more negative in this mutant, suggesting allosteric modulation promotes increased flexibility for this mutant enzyme. This is counterintuitive and provides clues to the anomalous exaggerated behaviour of this mutant. Binding of TIH to Δ258A saw the largest change to $\Delta C_p^{\ddagger}$ across each mutation and allosteric state, also accompanied by a further reduced $\Delta S^{\ddagger}$, not completely offset by the alteration to $\Delta H^{\ddagger}$ which became negative due to the notable shift in T_{opt}. TIH also elicited a reduction to the T_{opt} of Δ258A in favour of increased rates. The allosteric activation properties of TIH were altered in Δ258A, disappearing at high temperatures as a result of the altered $\Delta C_p^{\ddagger}$, as temperature dependent $\Delta C_p^{\ddagger}$ alters the curvature of the system, driving each dataset to overlay at the high temperature range.

Though this data was not collected at a physiologically relevant pH range, the general trends observed in this work are able to show how allosteric modulation can be modified in response to temperature. Gaining efficient allosteric regulation at one temperature could be through changing $\Delta C_p^{\ddagger}$ and hence, the dynamics of the system (**Figure 4.12** and **Table 4.2**), however, this comes at the expense of regulation at temperature extremes.

4.4 Conclusion

*mtu*ATP-PRT$_{WT}$ shows temperature independent reactions, and allosteric modulation by L-His and TIH, with alterations to the $\Delta C_p^{\ddagger}$ of the enzyme that affect allosteric modulation at upper and lower temperatures. We have also identified the role of residue 258 aspartic acid in the communication of allosteric inhibition signal from the C-terminal allosteric site to the catalytic site. Δ258N with alterations to the charge of the sidechain became insensitive to allosteric inhibition, however, activation by TIH remained suggesting this residue acts in allosteric inhibition alone. MMRT analysis shows an enthalpy-entropy trade-off that alters between allosteric states in *mtu*ATP-PRT. This trade-off reflects both non-covalent bond changes

between APO and L-His bound and the relative flexibility of these states that alter the energy barrier of the reaction, and thus in the case of L-His inhibition, decrease the reaction rate.

Chapter 5
Structural characteristics of allosteric regulation in *mtu*ATP-PRT

Crystal structures of *mtu*ATP-PRT$_{WT}$ have been solved in APO, L-His, and TIH bound states previously (Cho *et al.*, 2003; Pisco *et al.*, 2017). These structures compare closely with other ATP-PRT$_L$ structures and show the conformational changes between APO and L-His bound structures. These changes tighten the overall structure of the enzyme, causing a predicted overlap of substrate binding sites, and thus altering catalysis (Jiao *et al.*, 2019). Structures by Pisco *et al.* (2017) were the first to show that both L-His and TIH bind into the same allosteric binding pocket and elicit the same tensed conformation in the overall enzyme structure. These overlapping structural results confound the idea of allosteric inhibition being a consequence of tensed enzyme states, suggesting allosteric inhibition may not be driven by overlapping substrate binding sites. These structural changes do not agree with the differences in catalytic effect upon the binding of L-His and TIH, and they do not elucidate key residues or movements associated with allosteric modulation of *mtu*ATP-PRT. Site-directed mutagenesis reported here (Section 4.2.1) at position D258 ΔA and ΔN below the allosteric binding site resulted in changes to activity and sensitivity to histidine. Allosteric activation by TIH appeared unaffected by mutation (**Figure 4.12**) although inhibition and activation appear to be enhanced at low temperature in the *mtu*ATP-PRT$_{D258A}$ mutation. From this intriguing kinetic data, crystallography presents a further means to investigate these mutants and effector systems.

To determine how modification at position D258 alters allosteric regulation and side chain interactions that may contribute to the communication pathways of allosteric activation and inhibition, crystal structures of *mtu*ATP-PRT$_{D258A}$ and *mtu*ATP-PRT$_{D258N}$ were solved in APO, L-His, and TIH bound structures. These structures were used for careful analysis of the overall structure, RMSD, side chain interactions, and hydrogen bonding to investigate interactions within the molecule that may contribute to allosteric effects. In addition, these structures can be used as

starting structures for molecular dynamics simulations to investigate the role that protein dynamics may play in allosteric regulation.

5.1 Crystallography methods

For both *mtu*ATP-PRT variants with the various allosteric modulators, the hanging drop method combined with crystallisation conditions used by Pisco *et al.* (2017) and Cho *et al.* (2003) were used, with small variations in pH and precipitant to optimise crystal growth (Section 3.8.1). Crystals were looped, cryoprotected in mother liquor (with the addition of 10 % v/v glycerol) and flash frozen in liquid N_2. No correlation between morphology and diffraction data was observed. Flash frozen crystals were sent to the Australian Synchrotron and data were collected using varying attenuation and detector distances as described in **Table 5.1**. Collected data were indexed, integrated, and scaled using XDS (Kabsch, 2010) and each space group was determined with Pointless (Evans, 2011). Datasets were solved and refined using multiple programme suites in CPP4 and Phenix as described in (Section 3.8.3). PyMOL symmetry operators were used to generate the hexameric assembly of *mtu*ATP-PRT for further investigation. The final crystal growth conditions are summarised in **Table 5.1**.

5.2 Results

Crystals from each mutant in APO, L-His, and TIH grew in similar crystallisation conditions, and were solved between 2.0 – 2.5 Å resolution, making them ideal for comparison. All *mtu*ATP-PRT datasets belonged to the H 3 2 space group, and all contained a single protein chain within the asymmetric unit.

5.2.1 *mtu*ATP-PRT$_{\Delta 258}$ crystallisation

Crystals grew in two major morphologies; triangular prisms or cubes and to various sizes (**Figure 5.1**, panel **A** and **C** respectively). Triangular prisms for *mtu*ATP-PRT$_{WT}$ have been previously described in the literature (Moggré *et al.*, 2017). Cubic shaped crystals occurred only in conditions containing allosteric ligands L-His or TIH. Crystals successfully co-crystallised with allosteric ligands L-His and TIH

tended to adopt the cubic morphology, compared to APO crystals which formed triangular prisms.

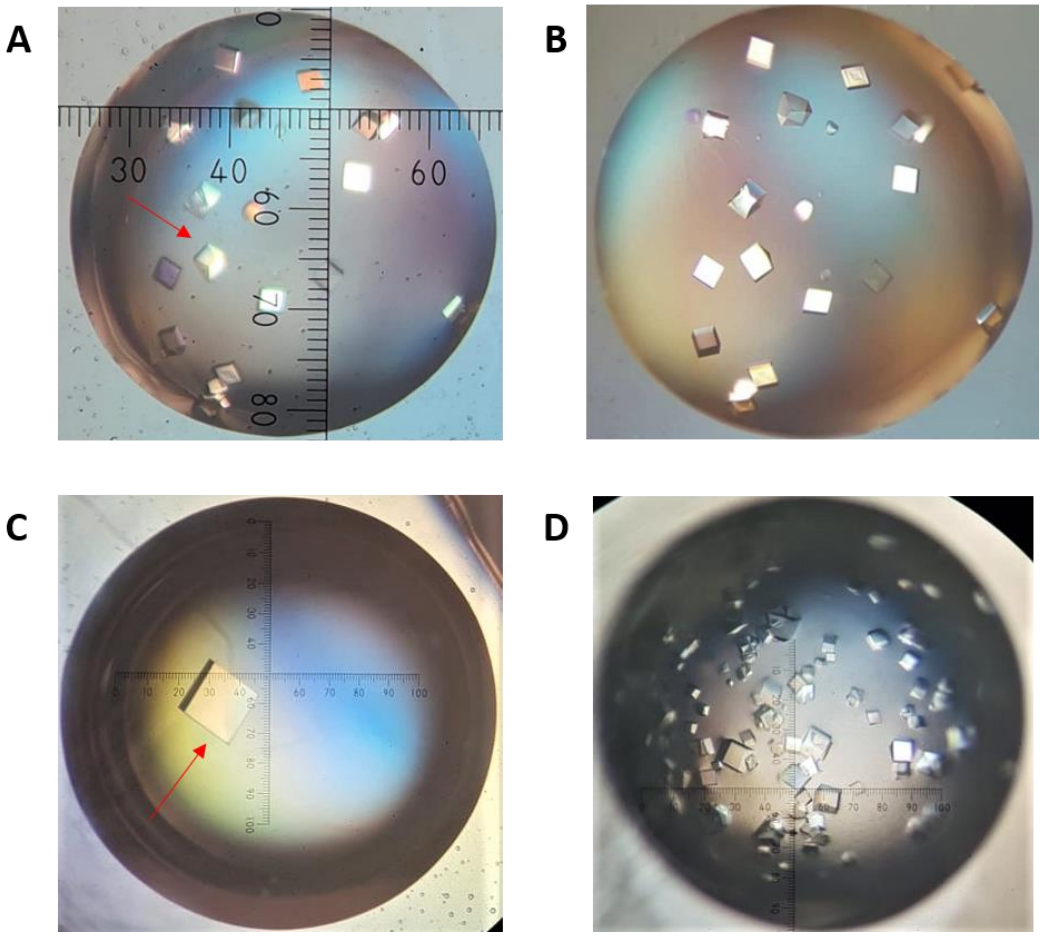

Figure 5.1. Crystal morphology of *mtu*ATP-PRT variants in hanging drop fine screens. **A**. *mtu*ATP-PRT$_{D258A}$ co-crystallised with TIH. **B**. mtuATP-PRT$_{D258A}$ co-crystallised with TIH. **C**. *mtu*ATP-PRT$_{D258N}$ co-crystallised with TIH. D. *mtu*ATP-PRT$_{D258N}$ co-crystallised with L-His. Panel **A** shows an arrow pointing to the previously described triangular prism type morphology, panel **C** points towards cubic crystal morphology.

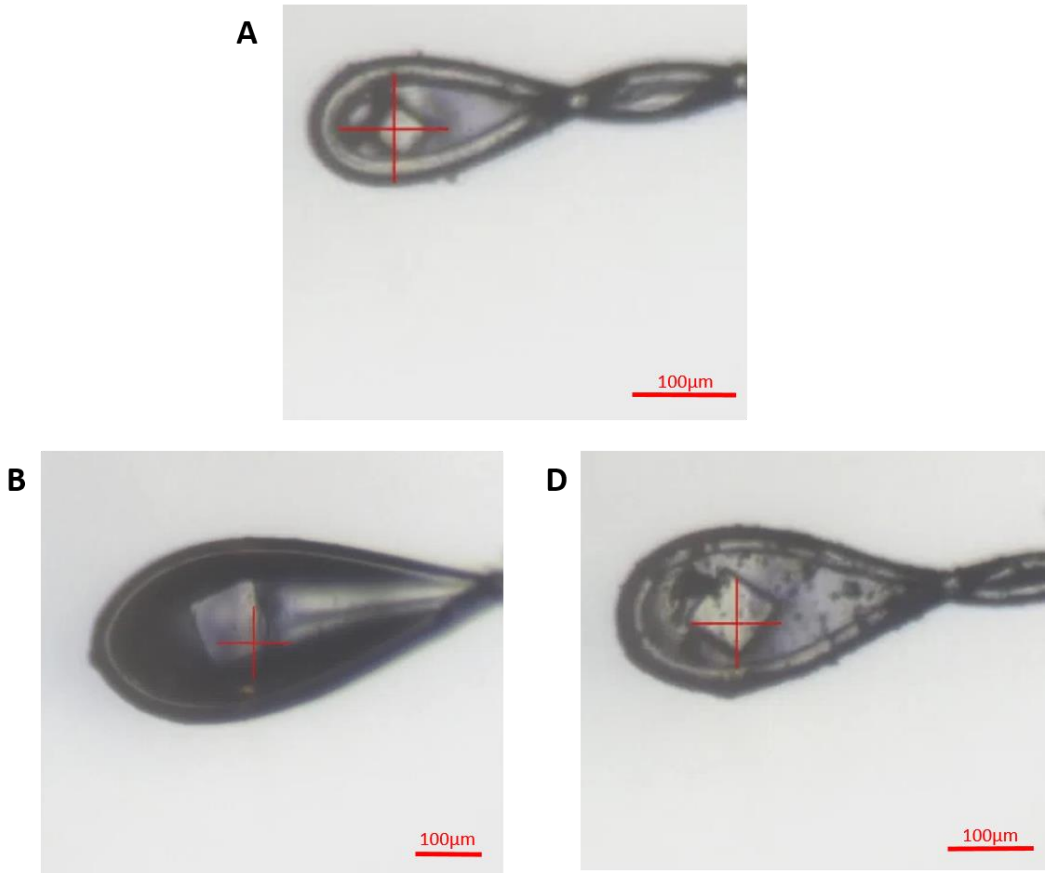

Figure 5.2. Crystals on loops for data collection at Australian Synchrotron MX2. Figures **A. B.** and **C**. represent the range of crystal morphologies and sizes of *mtu*ATP-PRT$_{D258A}$. These represent the range of crystals observed across all mutants.

Table 5.1. Crystallisation and data collection conditions for the final datasets. Concentration of purified *mtu*ATP-PRT was between 9 – 25 mg/mL.

Variant	Ligand	Crystallisation conditions	Detector distance (mm)	Attenuation (%)	Collection time (seconds)
D258A	APO	1.2 M $MgSO_4$, 0.1M MES (pH 6.5)	325	40	36
	L-His	1.0 M $(NH_4)_2SO_4$, 2.5 % IPA, 200 µM L-His	300	60	36
	TIH	1.0 M $MgSO_4$, 0.1 M MES (pH 6.1), 5.5 mM TIH	275	80	72
D258N	APO	1.6 M $MgSO_4$, 0.1 M MES (pH 6.7)	275	80	72
	L-His	1.2 $(NH_4)_2SO_4$,2.5% IPA, 200 µM L-His	280	70	36
	TIH	1.4 M $MgSO_4$, 0.1 M MES (pH 6.1), 5.5 mM TIH	325	25	36

5.2.2 Data collection statistics

Data from section 3.8.2 is described in **Table 5.2** and **Table 5.3** generated from Phenix (Liebschner *et al.*, 2019).

Table 5.2. Data collection statistics for *mtu*ATP-PRT$_{D258A}$ APO, L-His, and TIH bound structures.

Data statistic	***mtu*ATP-PRT$_{D258A}$ APO**		***mtu*ATP-PRT$_{D258A}$ with L-His**		***mtu*ATP-PRT$_{D258A}$ with TIH**	
Space group	H 3 2		H 3 2		H 3 2	
Resolution range (Å)	37.02 - 2.14		46.13 - 2.50		42.07 - 2.00	
	(2.22 - 2.14)		(2.59 - 2.50)		(2.07 - 2.00)	
Unit cell						
a/b/c (Å)	132.40/132.40/		114.56/114.56/125.43		114.96/114.96/126.202	
α/β/γ (°)	111.05		90/90/120		90/90/120	
	90/90/120					
Total reflections	420451	(41514)	229327	(23201)	437877	(45555)
Unique reflections	20628	(2033)	11147	(1081)	21861	(2185)
Multiplicity	20.4	(20.4)	20.6	(21.5)	20.0	(20.8)
Completeness (%)	99.78	(98.78)	99.96	(100.00)	99.96	(100.00)
Mean I/sigma(I)	36.56	(3.15)	25.47	(2.32)	24.34	(2.64)
Wilson B-factor	51.60		68.61		41.27	
R-merge	0.05	(0.90)	0.08	(1.29)	0.07	(1.07)
CC $_{1/2}$	1	(0.93)	1	(0.85)	1	(0.87)
Reflections used in refinement	20615	(2032)	11145	(1081)	21856	(2185)
Reflections used for R-free[†]	1039	(101)	579	(64)	1089	(125)

**Statistics for the highest-resolution shell are shown in parentheses. [†]5 % of reflections set aside for R-Free.*

Table 5.3. Data collection statistics for *mtu*ATP-PRT$_{D258N}$ APO, L-His, and TIH bound structures.

Data statistic	**mtuATP-PRT$_{D258N}$ APO**		**mtuATP-PRT$_{D258N}$ with L-His**		**mtuATP-PRT$_{D258N}$ with TIH**	
Space group	H 3 2		H 3 2		H 3 2	
Resolution range(Å)	40.35 - 2.2		46.41 - 2.3		41.84 - 2.15	
	(2.28 - 2.2)		(2.3882 - 2.3)		(2.23 - 2.153)	
Unit cell						
a/b/c (Å)	132.28/132.28/111.261		115.29/115.29/125.88		114.71/114.71/125.53	
α/β/γ (°)	90/90/120		90/90/120		90/90/120	
Total reflections	393986	(38186)	299824	(29364)	350496	(34156)
Unique reflections	19134	(1899)	14507	(1461)	17396	(1711)
Multiplicity	20.6	(20.1)	20.7	(20.1)	20.1	(20.0)
Completeness (%)	99.91	(99.89)	99.95	(100.00)	99.87	(99.13)
Mean I/sigma(I)	32.06	(2.64)	34.11	(3.81)	23.88	(2.90)
Wilson B-factor	51.83		53.39		48.72	
R-merge	0.06222	(1.198)	0.06008	(0.84)	0.07389	(0.9771)
CC $_{1/2}$	1	(0.92)	1	(0.90)	1	(0.91)
Reflections used in refinement	19121	(1897)	14503	(1461)	17393	(1711)
Reflections used for R-free†	952	(89)	672	(69)	880	(77)

**Statistics for the highest-resolution shell are shown in parentheses. †5 % of reflections set aside for R-Free.*

5.2.3 Data refinement statistics

Datasets were visualised and built manually in Coot (Emsley *et al.*, 2010) and refined using Refmac5 (Murshudov *et al.*, 1997) until each structure was completed as far as the density allowed. Additional density was fitted with water and metal ions, where chemistry and density were supported. Refinement statistics tables (**Table 5.4** and **Table 5.5**) were generated in Phenix (Liebschner *et al.*, 2019).

Table 5.4. Refinement and model statistics for *mtu*ATP-PRT$_{D258A}$ in APO, L-His, and TIH bound forms.

Data statistic	***mtu*ATP-PRT$_{D258A}$ APO**	***mtu*ATP-PRT$_{D258A}$ with L-His**	***mtu*ATP-PRT$_{D258A}$ with TIH**
R-work	0.22 (0.30)	0.20 (0.29)	0.20 (0.25)
R-free	0.27 (0.35)	0.26 (0.34)	0.26 (0.31)
Number of non-hydrogen atoms	2139	2075	2182
Ligands	27	21	23
Solvent	43	9	73
Protein residues	277	278	278
RMS(bonds)	0.014	0.014	0.015
RMS(angles)	1.90	2.07	1.90
Ramachandran favoured, allowed, disallowed (%)	98.15, 1.85, 0.00	94.53, 3.65, 1.82	95.24, 4.40, 0.37

**Statistics for the highest-resolution shell are shown in parentheses.*

Table 5.5. Refinement and model statistics for *mtu*ATP-PRT$_{D258N}$ in APO, L-His, and TIH bound forms.

Data statistic	**mtuATP-PRT$_{D258N}$ APO**	**mtuATP-PRT$_{D258N}$ with L-His**	**mtuATP-PRT$_{D258N}$ with TIH**
R-work	0.21 (0.29)	0.20 (0.22)	0.21 (0.25)
R-free	0.25 (0.33)	0.25 (0.27)	0.24 (0.31)
Number of non-hydrogen atoms	2106	2113	2195
Ligands	36	25	10
Solvent	31	26	34
Protein residues	276	284	286
RMS(bonds)	0.01	0.01	0.02
RMS(angles)	1.96	1.90	1.92
Ramachandran favoured, allowed, disallowed (%)	95.96, 4.04, 0.00	96.77, 2.51, 0.72	96.11, 3.89, 0.00

**Statistics for the highest-resolution shell are shown in parentheses.*

5.2.4 Identification of allosteric site ligand

Due to the dramatic change to the overall structure upon allosteric ligand binding, identification of successful co-crystallisation was determined based on log likelihood gain (LLG) values produced in Phenix Phaser. Structures containing ligand, and thus fit to the correct "tensed" model resulted in higher LLG values. To further confirm the identity of the ligand careful attention was paid to the density surrounding the ligand during refinement. Due to the extra density of the sulphur in the ring of TIH, positive and negative density could be observed during refinement that aided in the orientation of the molecule (**Figure 5.3**). Density surrounding the ligands in *mtu*ATP-PRT$_{D258A}$ and *mtu*ATP-PRT$_{D258N}$ is shown as omit maps in **Figure 5.4**.

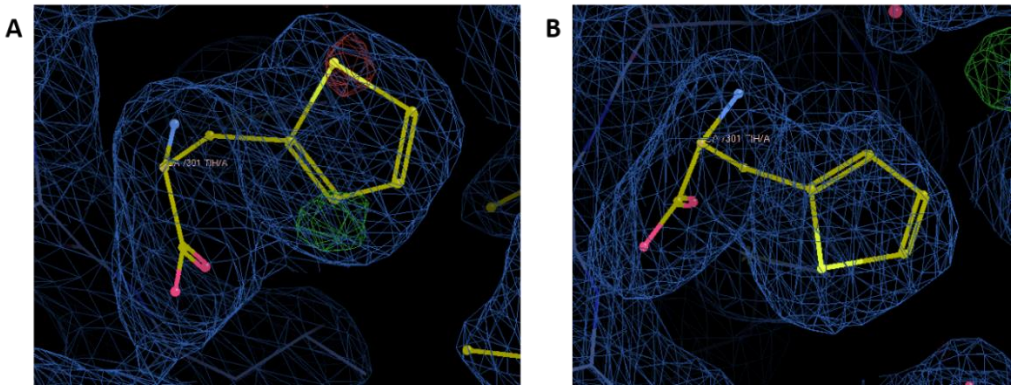

Figure 5.3. *mtu*ATP-PRT$_{D258N}$ ligand density at the trimeric interface. Panel **A** has TIH fit into the density and following refinement shows excess density in the upper region with lacking density in the lower direction. Panel **B** shows TIH correctly oriented in the density, satisfying density after refinement. From this, confident orientation of the ligand can be achieved.

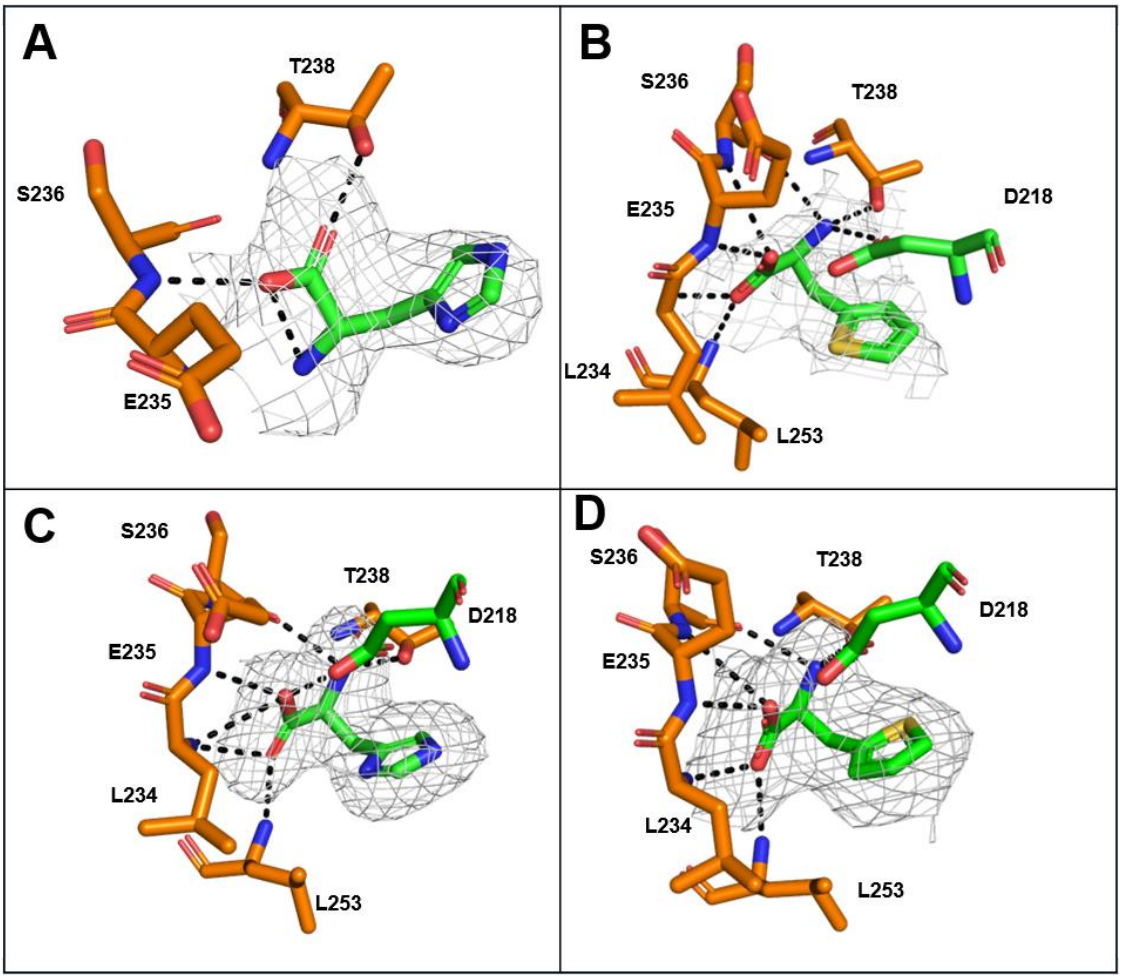

Figure 5.4. Omit map surrounding allosteric ligands in each *mtu*ATP-PRT variant generated in Phenix Polder (Liebschner *et al.*, 2017). **A**. *mtu*ATP-PRT$_{D258A}$ co-crystallised with L-His. **B**. *mtu*ATP-PRT$_{D258A}$ co-crystallised with TIH. **C**. *mtu*ATP-PRT$_{D258N}$ co-crystallised with L-His. **D**. *mtu*ATP-PRT$_{D258N}$ co-crystallised with TIH.

5.2.5 Overall structure

The overall structure of each variant was consistent with the wildtype structures (PDB: 1NH7, 5LHU, 5LHT (Cho *et al.*, 2003; Pisco *et al.*, 2017)), as expected given the single amino acid variations between them. Each dataset contained one monomer per asymmetric unit, the hexamer representing biological assembly (homohexamer) were obtained using symmetry operations in PyMOL. The completed monomer and hexameric assembly of *mtu*ATP-PRT$_{D258A}$ is illustrated in **Figure 5.5**, and the topology of *mtu*ATP-PRT$_{D258N}$ is shown in **Figure 5.6**. These structures are not vastly different to other published *mtu*ATP-PRT structures (Cho *et al.*, 2003; Moggré *et al.*, 2017; Pisco *et al.*, 2017), therefore they are briefly described, and focus is placed later on in this thesis to describe fine details between each structure.

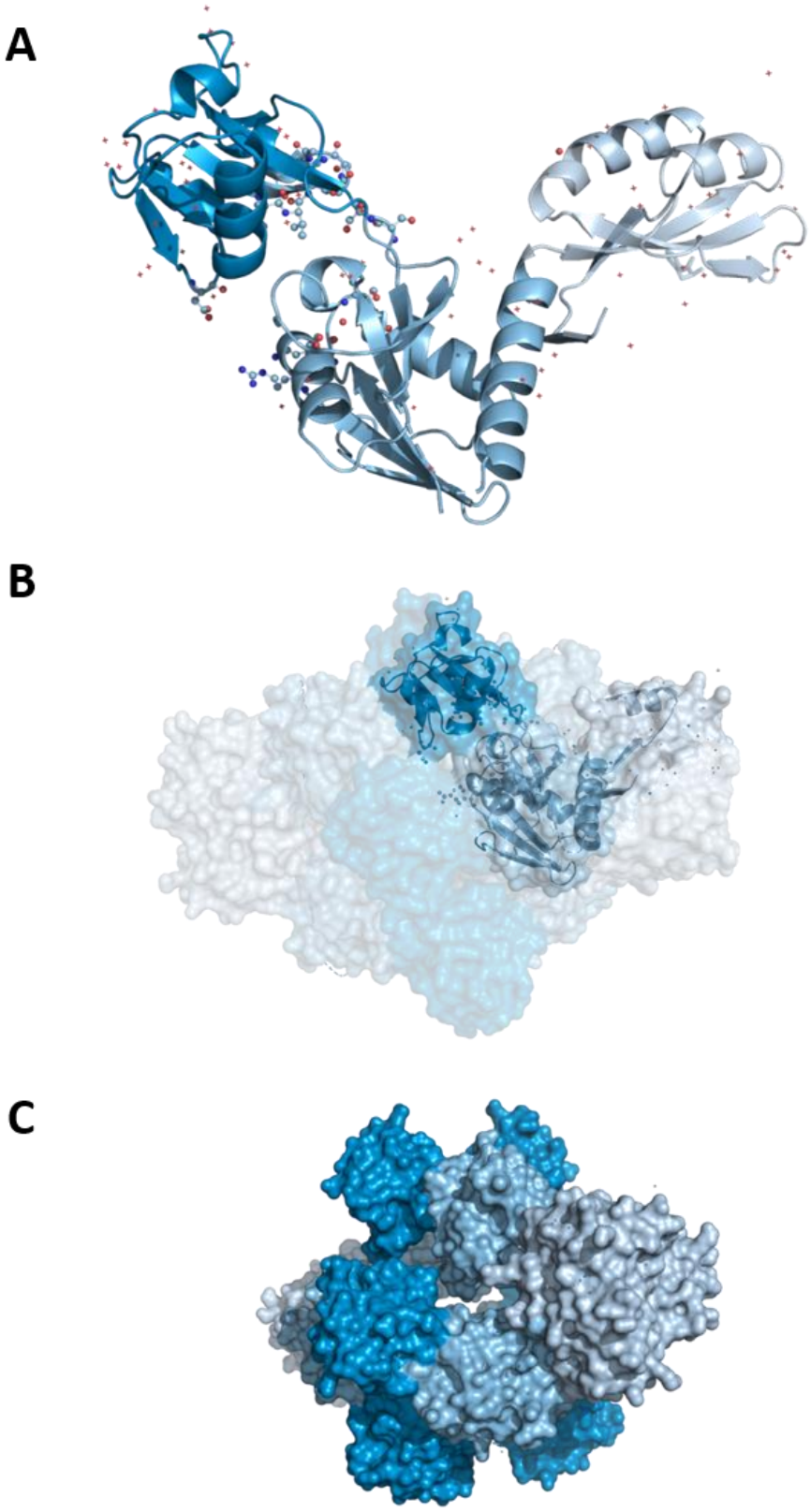

Figure 5.5. 3D structure of *mtu*ATP-PRT$_{D258A}$ with TIH bound. The monomer (**A**) forms a dimer-trimer (**B**) assembly forming a homohexamer (**C**). Protein surface displays large gaps through the centre of the hexamer assembly when in the tensed state (**C**). Domains coloured; domain I (medium blue), domain II (dark blue), domain III (pale blue). Light blue colouring is carried through the chapter for Δ258A with TIH bound. TIH is shown as sticks, and predicted active site residues are represented as ball and stick in panel **A**.

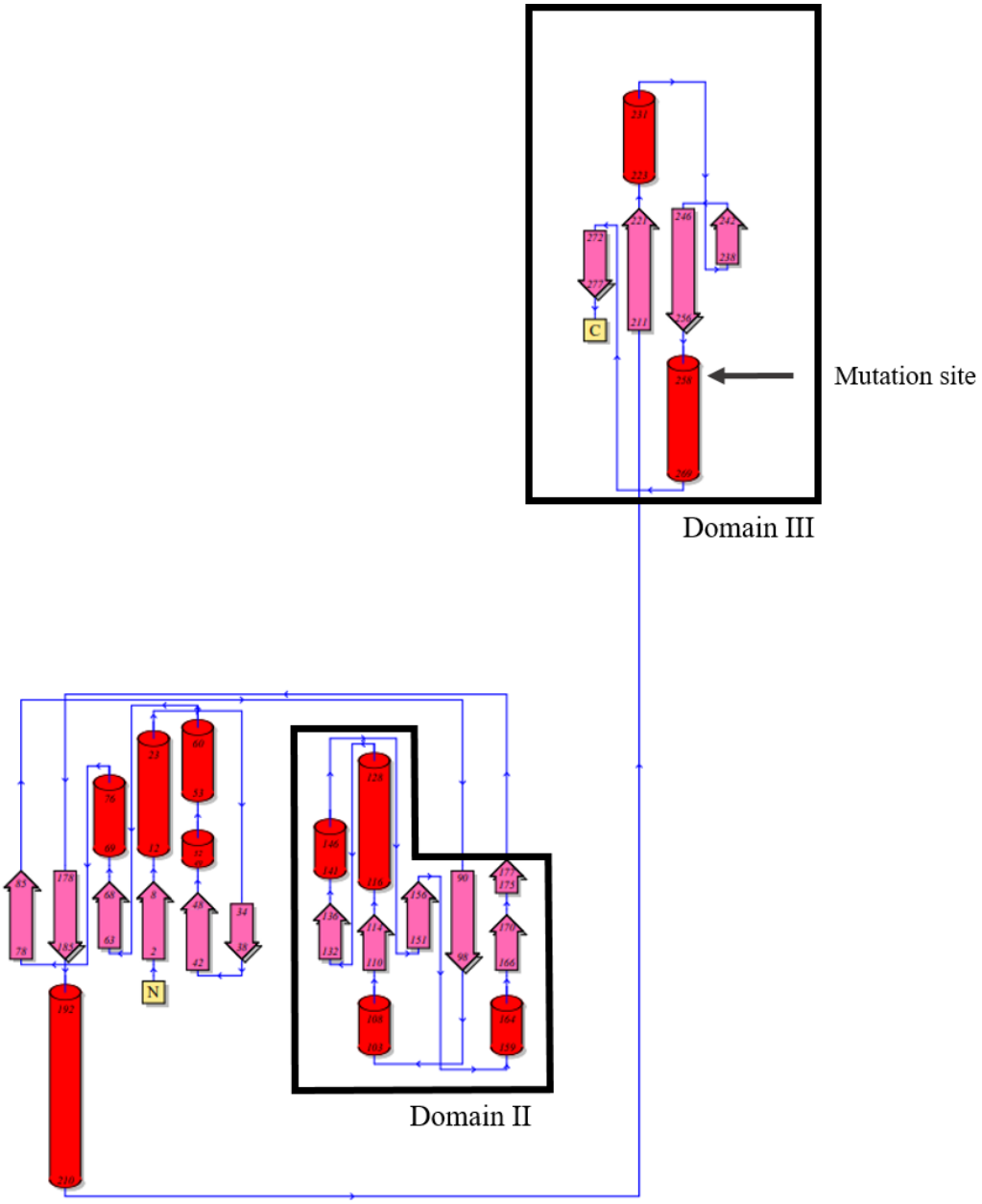

Figure 5.6. Topology structure of *mtu*ATP-PRT. Sheets (pink) and helices (red). Domains II and III have been boxed. The βαββαβ structure of Domain III is the location of the allosteric binding site and site mutation Δ258. Figure made in PDBsum (Laskowski *et al.*, 1997).

Each of the six solved structures are missing a small section of a loop near the eighth helix leading up toward domain III (**Figure 5.7**). *mtu*ATP-PRT$_{D258A}$ (APO) had an additional missing residue in a tight turn section leading into a β-sheet of the regulatory domain. These missing residues are unconstrained loops on the surface

of the enzyme. This is a common feature in crystallographic data and is seen in other ATP-PRT structures (PDB: 1NH7 (Cho *et al.*, 2003)).

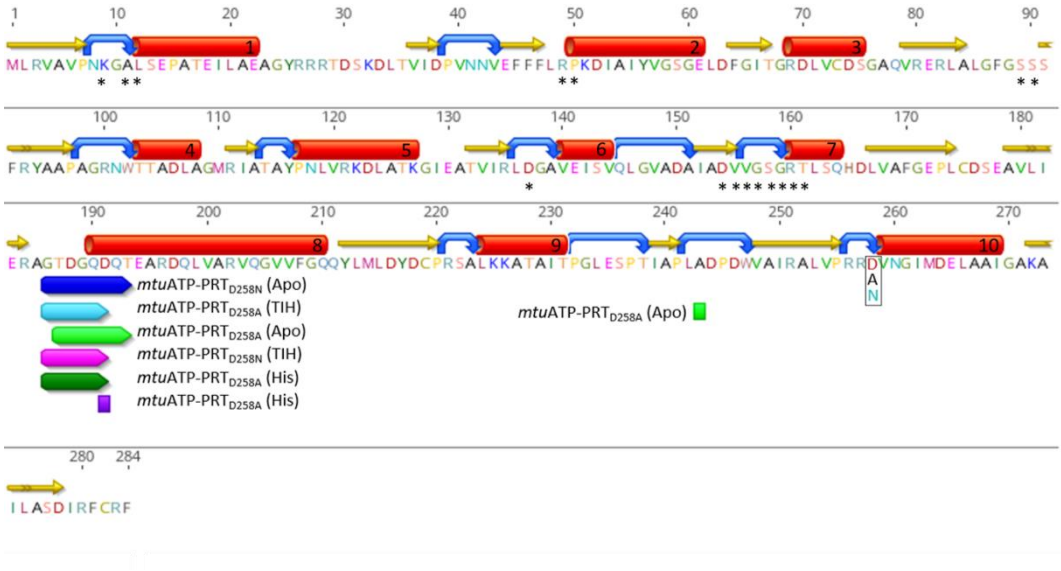

Figure 5.7. Missing residues for each solved structure as indicated by coloured bars. Active site residues are starred. The site of mutation is boxed. Each structure is missing 1 – 7 residues in the solvent exposed loop section leading into the 8th helix. In addition, *mtu*ATP-PRT$_{D285A}$ in APO form is missing an additional residue (244) in a tight turn in Domain III. Secondary structure annotations were added in Geneious Prime (Geneious Prime 2019.2.1 (https://www.geneious.com)). Helices are represented by red boxes, turns represented by blue arrows, and sheets represented as yellow arrows.

5.2.5.1 Ligands and disulphide bonds

In addition to allosteric ligands (when present), glycerol and sulphate were found across all structures. Both glycerol and sulphate were in abundance in crystallographic and cryoprotectant conditions. Residues C73 and C175 appeared to form disulphide interactions in some structures. This disulphide interaction is located between domain I and II of the active site (**Figure 5.8**), with density to support the presence of alternating conformations in some structures (**Figure 5.9** and **Figure 5.10**). For *mtu*ATP-PRT$_{D258A}$ APO and L-His bound, this disulphide interaction is present with no density to support alternative conformations (**Figure 5.9 A** and **B**). However, with TIH bound, density around C73 and C175 supported alternate conformations for these residues, potentially indicating both a bridged and unbridged form (**Figure 5.9 C**). In *mtu*ATP-PRT$_{D258N}$, more variation is seen in the presence and absence of this disulphide interaction. This interaction is absent in the APO structure (**Figure 5.10 A**). In L-His bound, there is a section of negative

density across the bond, (**Figure 5.10 B**) however, there is no density to support the alternative conformation and therefore has been left in place as a disulphide bridge. Finally in **Figure 5.10 C** we see alternate conformations of the two cysteine residues, the same as seen in the other TIH bound structure (**Figure 5.9 C**). Both conformations have been observed in other *mtu*ATP-PRT$_{WT}$ structures; *e.g.* a disulphide is present in 1NH7, 5LHU, and 5LHT (Cho *et al.*, 2003; Pisco *et al.*, 2017), and absent in 5U99 (Moggré *et al.*, 2017). Reducing agent (β-Mercaptoethanol) was only present at less than 1 mM during TEV protease digestion, and was not present in any purification or crystallisation conditions.

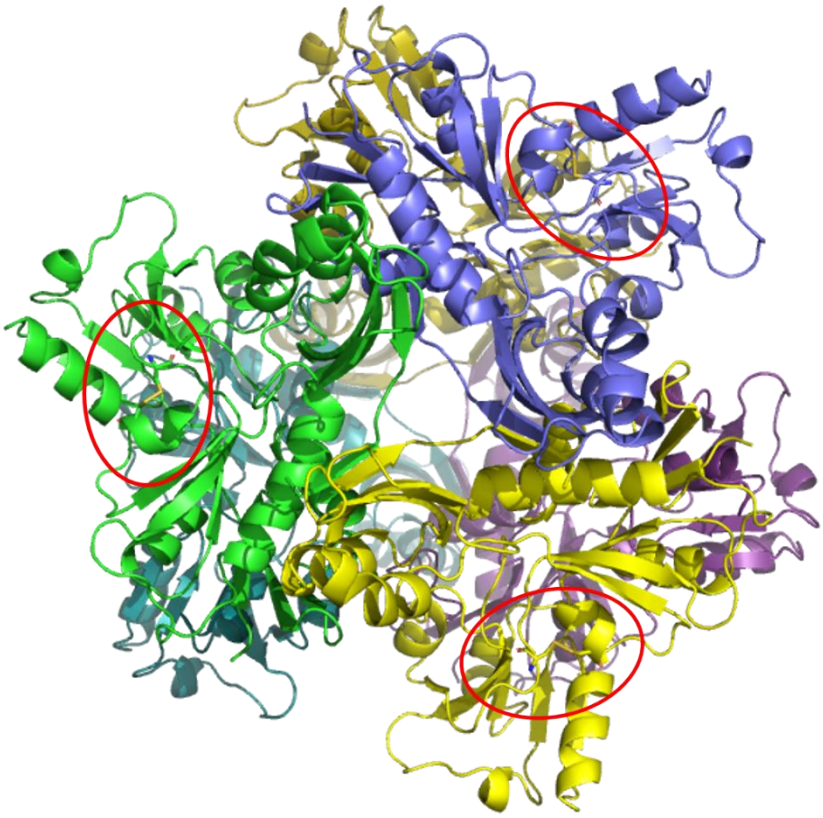

Figure 5.8. Cysteine bridge location in *mtu*ATP-PRT$_{D258A}$ (APO) structure (circled). When formed, this disulphide interaction is formed between Cys73 and Cys175 of domain I and II, near the active site pocket.

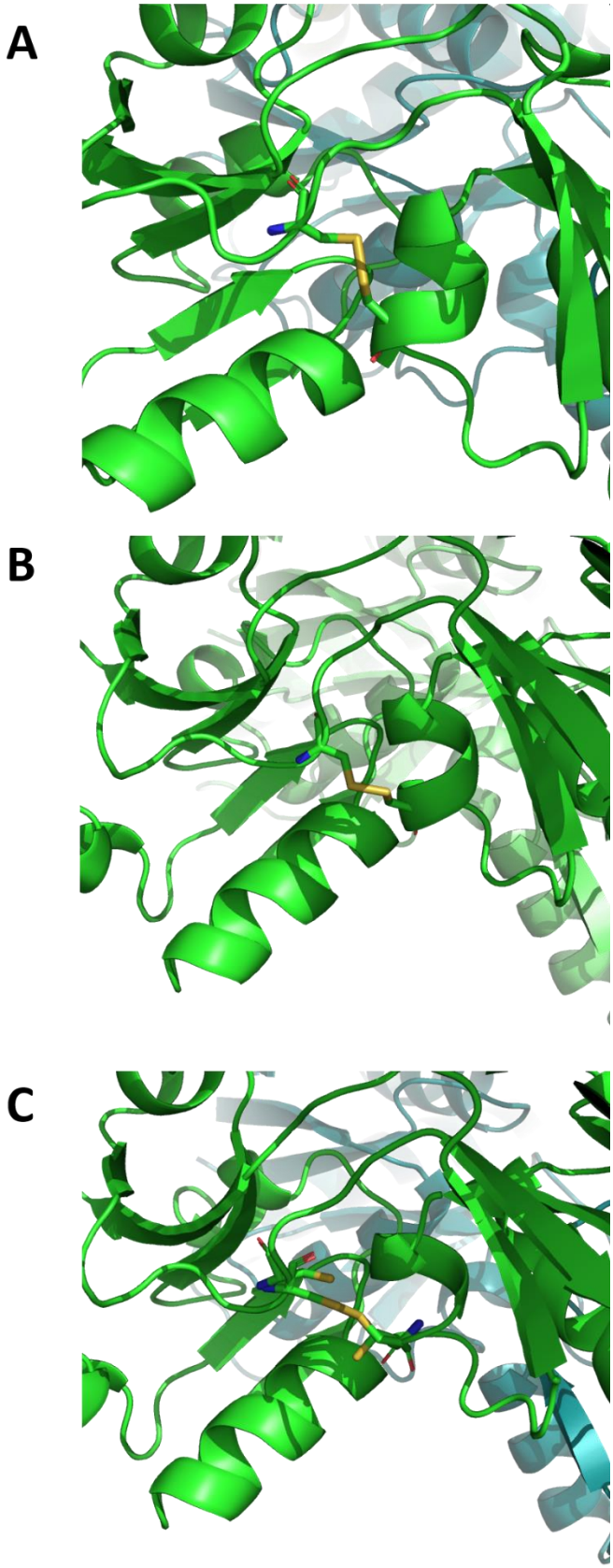

Figure 5.9. Cysteine residues 73 and 175 interaction in *mtu*ATP-PRT$_{D258A}$ crystal structures. **A.** Interaction in APO structure. **B.** Interaction of residues in L-His bound structure. **C.** Interaction of residues and alternate conformations in TIH bound structure.

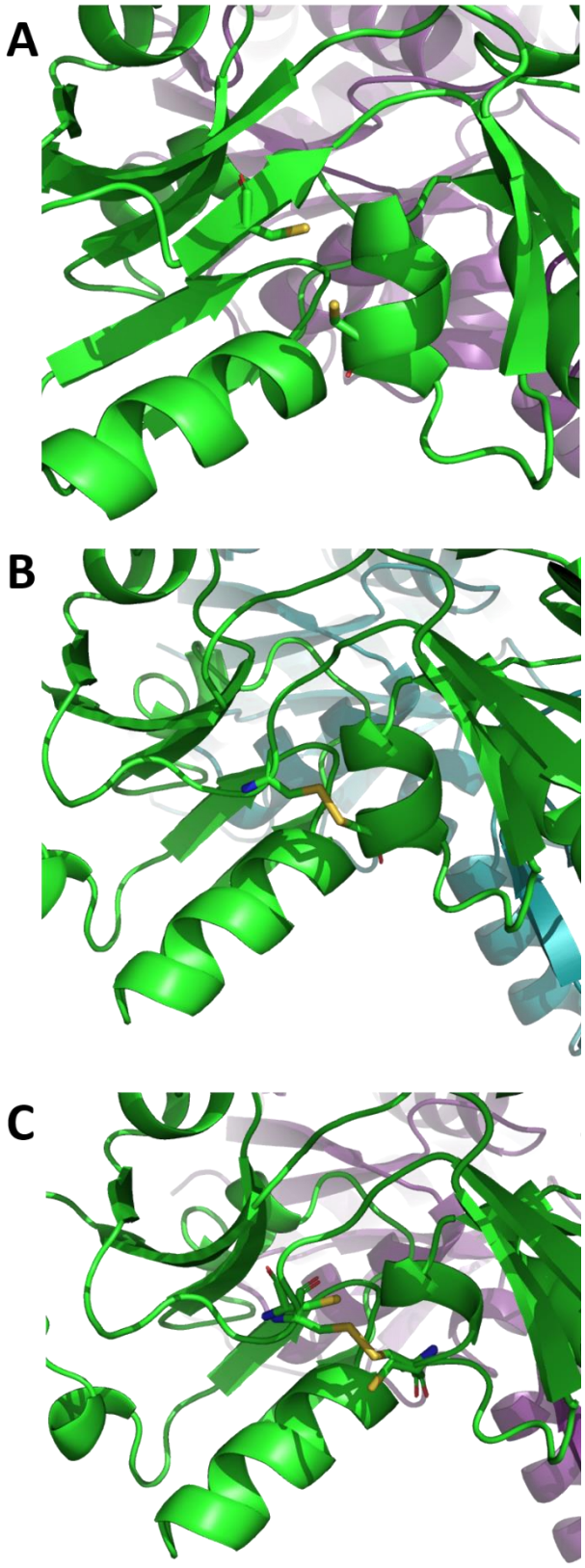

Figure 5.10 Cysteine residues 73 and 175 interaction in *mtu*ATP-PRT$_{D258N}$ crystal structures. **A.** Interaction in APO structure. **B.** Interaction of residues in L-His bound structure. **C.** Interaction of residues and alternate conformations in TIH bound structure.

At the trimeric interface of all six structures was a symmetrical section of unmodelled density which formed a trio of peaks around a three-fold axis (**Figure 5.11** panel **A**). In previously solved structures (PDB; 1NH7) this section of density had been fitted with magnesium (Cho *et al.*, 2003). Ca, Co^{2+}, and Mg^{2+} (possibly present from buffer and crystallisation conditions) were placed and refined into this section shown in **Figure 5.11** panel **B**, **C**, and **D**. As none of these metals explained the density, the difficulties of data accuracy at the three-fold axis, and the proximity of this position away from areas of interest in the structure, this position was left empty in the final structures. A wavelength scan was undertaken to try to identify the metal ion here and produced inconclusive results.

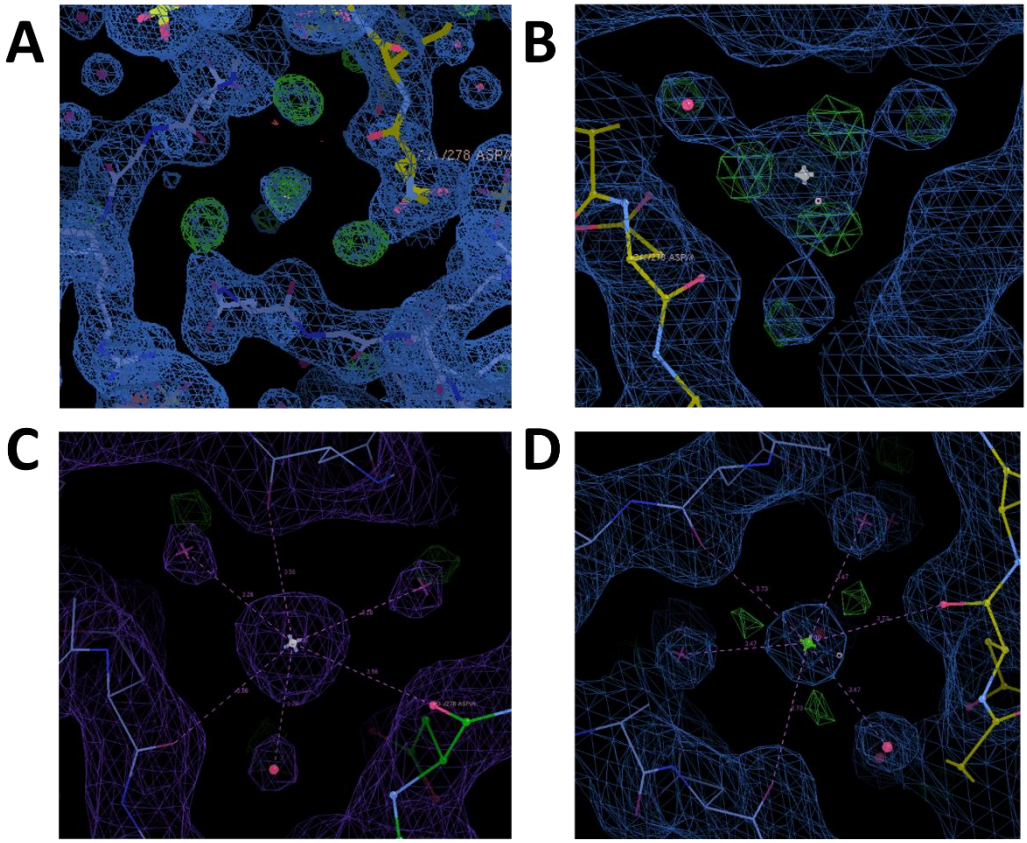

Figure 5.11. Density at the three-fold axis of solved crystal structures. **A.** Unmodelled density in the trimeric interface of *mtu*ATP-PRT variants, temporarily modelled with prospective metal ions. **B.** Calcium modelled into *mtu*ATP-PRT_{D258A} (APO). **C.** Co^{2+} - modelled into *mtu*ATP-PRTD258A (APO). **D.** Mg^{2+} modelled into *mtu*ATP-PRT_{D258N} (TIH). None of these metals explain the density, although the density is likely confounded from the close proximity to the three fold axis

5.2.6 Structure comparison

Following completion of the Δ258 mutant structures, completed structures were used alongside wildtype structures (PDB; 1NH7, 5LHU, 5LHT) to investigate changes to each structure upon mutation and allosteric ligand binding.

5.2.6.1 B-factor analysis

B-factors in proteins represent the uncertainty of residue positioning in a protein structure – either from structure flexibility or poor data quality. As the set of ATP-PRT structures are all of similar resolution, B factors present a reasonable way to assess relative protein flexibility differences across the data set. **Figure 5.12** graphs the average B-factor for each residue in the monomer. The large peaks correspond to some of the more flexible surface residues. There are missing residues in a loop between residues 184 – 193 in some datasets (**Figure 5.7**), with high B-factors either side, indicating these are highly flexible regions explaining the lack of density in missing residue. WT APO was removed from this analysis due to anomalous B-factors.

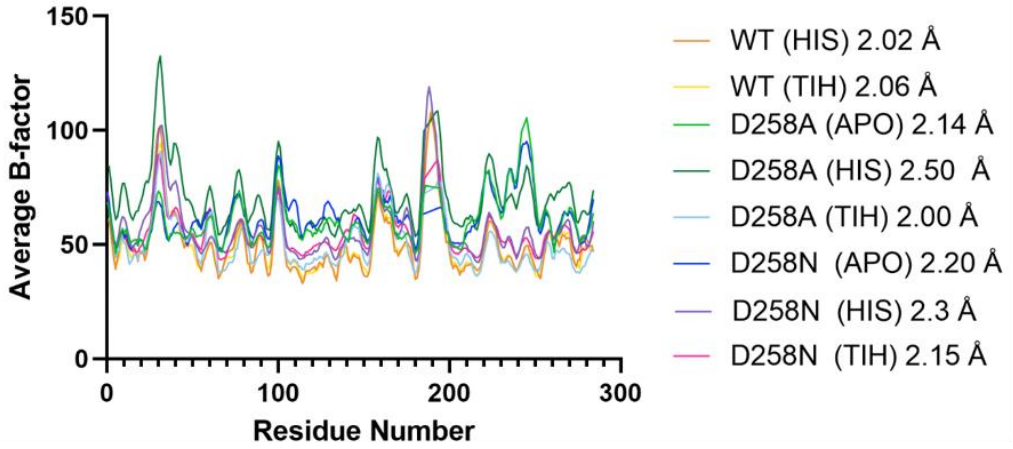

Figure 5.12. Average B-factor of each residue calculated in Baverage (Dodson, 1991). Coloured by protein, B-factor peaks signify increased local flexibility. Data has been smoothed in GraphPad Prism, gaps in data between residues 184 – 193 are a function of missing residues (outlined in **Figure 5.7**). Structure resolution is provided to gauge contributions to B-factor relative to data quality.

*mtu*ATP-PRT$_{D258A}$ (L-His) shows the highest average B-factors, particularly in the first 100 residues (domain I and II). *mtu*ATP-PRT$_{D258A}$ (L-His) has B-factors that remain high across the whole structure; however they are reduced between residues 110-140 and 230-260. Excluding these anomalous *mtu*ATP-PRT$_{D258A}$ (L-His) B-factors, the remaining ligand bound B-factors showed the lowest flexibility, most apparent between residues 105-150 and 200-260. B-factors across all structures aggregate around residues 170-195, which form a surface region in domain I and II.

Aside from *mtu*ATP-PRT$_{D258A}$ (L-His), the B-factors of L-His and TIH bound states of each variant seemingly track together across the molecule. In the *mtu*ATP-PRT$_{WT}$ structures there are no incidences where one ligand bound state elicits extreme differences to B-factor. In both *mtu*ATP-PRT$_{D258N}$ and *mtu*ATP-PRT$_{WT}$ L-His bound molecules tend to show slightly lowered B-factors for domain I and II, however this appears to switch around domain III (Residues 210 – end) in *mtu*ATP-PRT$_{D258N}$, where the L-His bound molecule sees slightly increased B-factor values. Additionally, in *mtu*ATP-PRT$_{D258N}$, there are three sections where the B-factor values diverge between L-His and TIH bound states. L-His binding elicits increased B-factors between residues 30-50 and 240-250, and decreased B-factors in between residues 140-150.

Figure 5.12 illustrates similar general trends in the B-factors of each protein variant. For the purpose of this thesis *mtu*ATP-PRT$_{D258A}$ (L-His bound) is displayed in **Figure 5.13**, however the B-factors of all variants were examined. Increased B-factors are seen in loop regions of each domain at chain interfaces (**Figure 5.13**). These surface loop regions are less constrained than internal residues of the protein and therefore have increased flexibility due to their position.

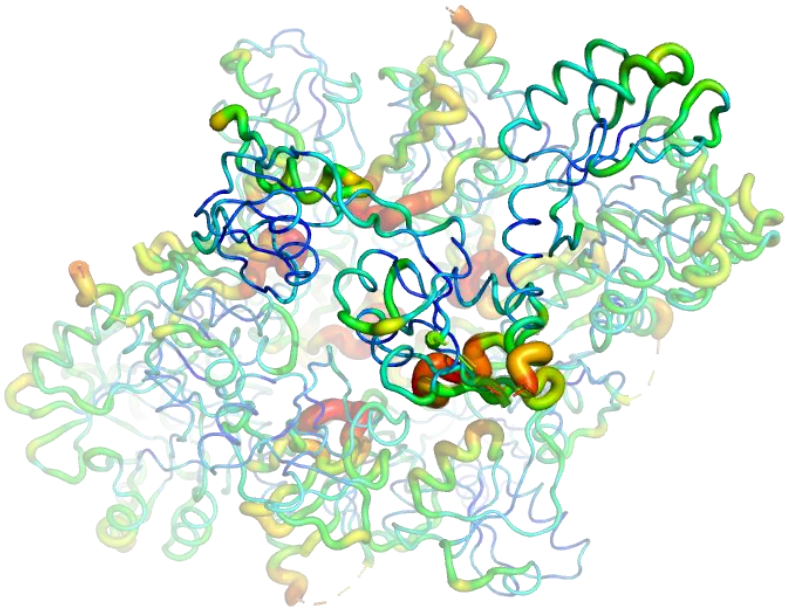

Figure 5.13. B-factor putty of the *mtu*ATP-PRT$_{D25A}$ (L-His bound) hexamer, illustrating increased B-factors at chain interfaces. Blue represents the lowest B-factor, and red represents the highest B-factors.

5.2.6.2 Root mean square deviation

Global root mean square deviation (RMSD) comparisons of the C_α atoms were undertaken to ascertain what structural movements differed between allosteric bound states, and if there were any changes in ATP-PRT variants that were non-responsive to allosteric inhibition.

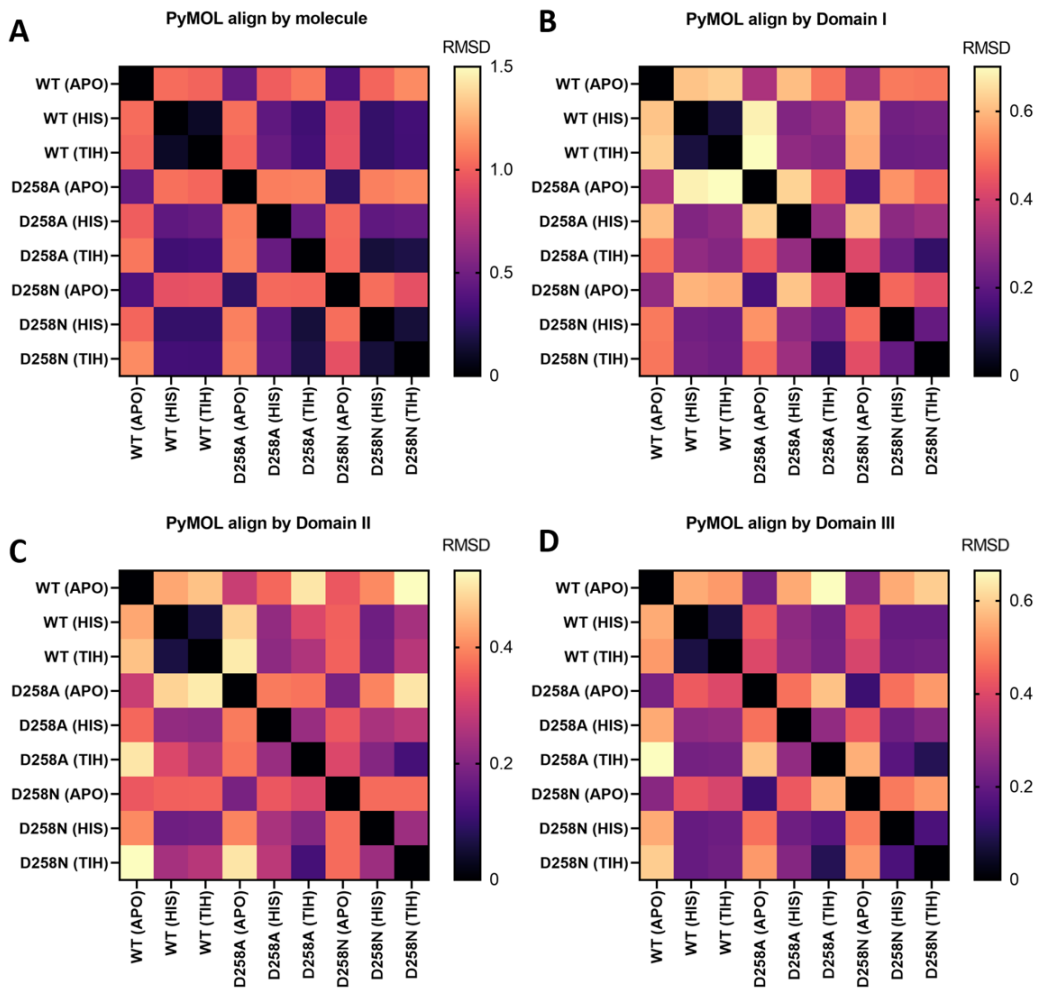

Figure 5.14. RMSD changes between allosteric bound states and ATP-PRT variants. RMSD values were collected when aligning whole molecule to whole molecule, and independent domains to independent domains to assess movement within each respective section. **A.** Global RMSD change of all three domains. **B.** RMSD change between Domain I (residue 1-89, 177-210). **C.** RMSD change between Domain II (resi 90-176). **D.** RMSD change between Domain III (resi 211-284). Note the changes in scale for each figure, with the largest variation in RMSD seen in A and the lowest in figure C (Domain II).

Figure 5.14 A shows larger RMSD changes across the global structure compared to figures **B**, **C**, and **D** indicating the global changes in domain orientation are the main alterations, compared to any changes within individual domains between allosteric states and protein variant. The average RMSD (Appendix 4, **Table A.4.1**) of the global protein was 0.602 compared to 0.352, 0.284, and 0.316 for Domain I, II, and III respectively. The particularly low mean RMSD for Domain II indicates little movement within the domain, and therefore possibly the least involvement in allosteric regulation. Across the global and domain specific alignments, *mtu*ATP-PRT$_{D258A}$ and *mtu*ATP-PRT$_{D258N}$ both show increased changes in RMSD between L-His and TIH bound states, demonstrating there is increased changes in

these structures compared to wildtype. This is particularly evident in each *mtu*ATP-PRT$_{D258A,}$ with each domain and the global RMSD change between L-His and TIH bound states being greater. *mtu*ATP-PRT$_{D258N}$ also shares increased RMSD in Domain II, similar to *mtu*ATP-PRT$_{D258A.}$

5.2.7 Structural changes surrounding Δ258

Residue 258, the mutation location, is located at the end of the second α helix of the regulatory C-terminal domain (domain III) (**Figure 5.15**). This residue is solvent exposed, and 7.5 Å and 8.0 Å from the nearest inter-chain residue in tensed and relaxed states respectively. Due to its position at the end of a helix and beginning of loop, position 258 can interact with its two flanking residues, Arg257 and Val259, and Pro255 that ends the third β-sheet of the regulatory domain. The distance between Asp258 and Pro255 decreases upon allosteric modulation by L-His from 3.9 Å to 3.4 Å (PDB 5LHU and 1NH7; (Cho *et al.*, 2003; Pisco *et al.*, 2017)) tightening the trimeric interface of domain III.

Figure 5.15. Location of mutation in WT (PDB 1NH7; (Cho *et al.*, 2003)). Asp258 interacts with Pro255 and moves closer tightening the loop during allosteric modulation into the tensed crystallographic state.

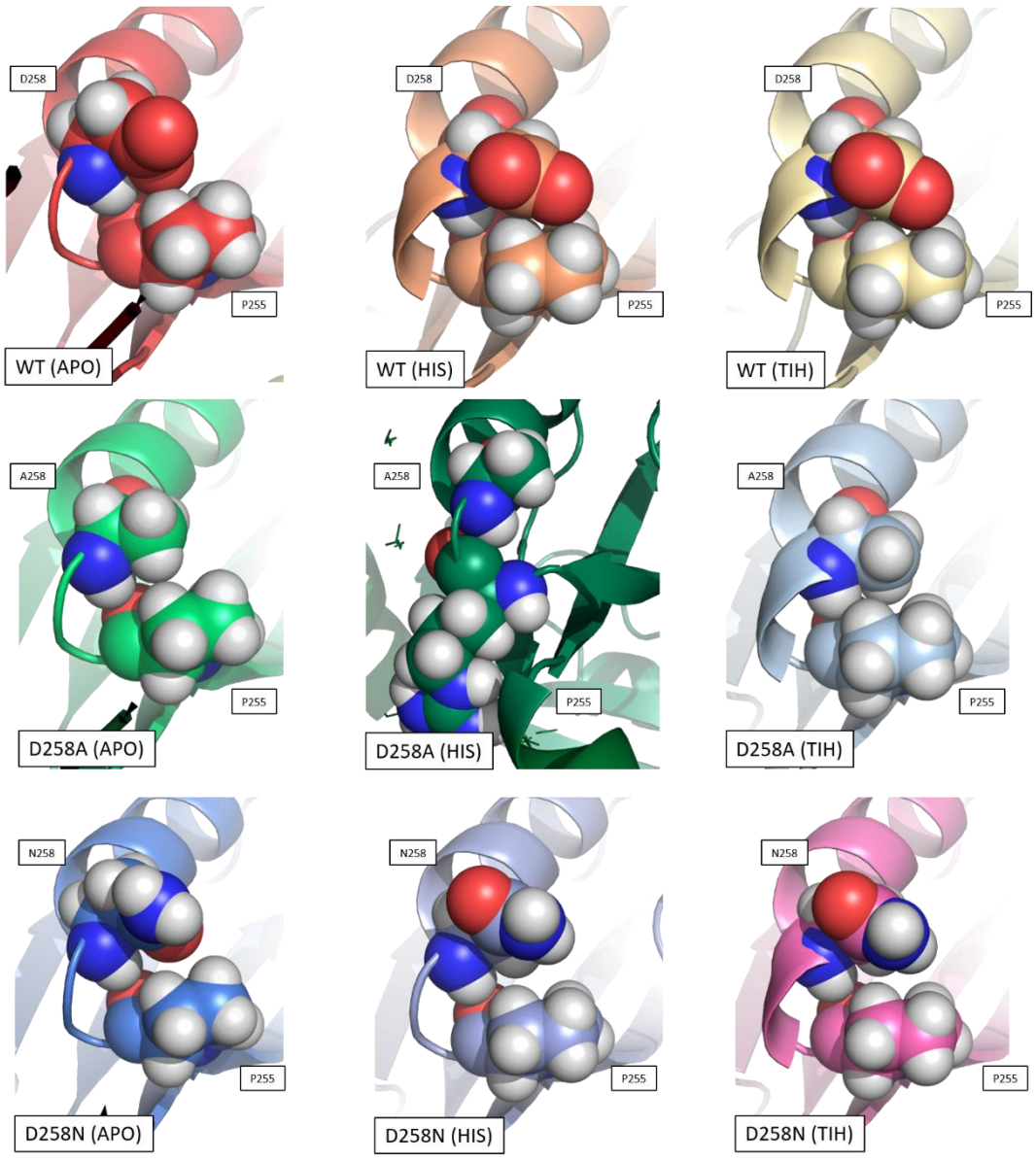

Figure 5.16. Interactions of residue 258D/A/N and Pro255 (spheres) across each *mtu*ATP-PRT variant. WT structures from PDB 1NH7, 5LHU, and 5LHT (Cho et al., 2003; Pisco et al., 2017).

In the allosterically modulated wild type structures Asp258 interacts tightly with Pro255 tightening the loop between the α-helix and β-sheet of the regulatory domain (domain III) (**Figure 5.16**). Where the side chain of Asp258 has been truncated to alanine or the charge removed in asparagine, this interaction appears to be altered. *mtu*ATP-PRT$_{D258A}$ APO and TIH bound structures show residue 258 to Pro255 interactions similar to wildtype. However, in the L-His bound structure Pro255 appears rotated and the interaction is removed, perhaps involved with anomalous kinetic data **(Figure 4.12**). In *mtu*ATP-PRT$_{D258N}$ the APO and L-His bound structures seem unaltered with a less close interaction (5.3 and 5.1 Å)

between the residues Asn258 and Pro255. Interestingly, in the TIH bound structure these two residues interact more closely (4.8 Å) compared to the APO structure although not as closely as the WT enzyme.

Jiao *et al.* (2019) observed in *C. jejuni* several residue contact changes on L-His binding. Arg216 interacts with Glu279 and Glu276 in APO state, and switches to an interaction with E271 in the allosterically inhibited state. The equivalent residues in *mtu*ATP-PRT are Arg202, Arg256, Asn260 and Asp264 (to *cje*ATP-PRT Arg216, Glu271, Glu276, and Glu279 respectively). These interactions are visualised in **Figure 5.17**, and interactions between residues in *mtu*ATP-PRT variants in each allosteric state (**Figure 5.18**).

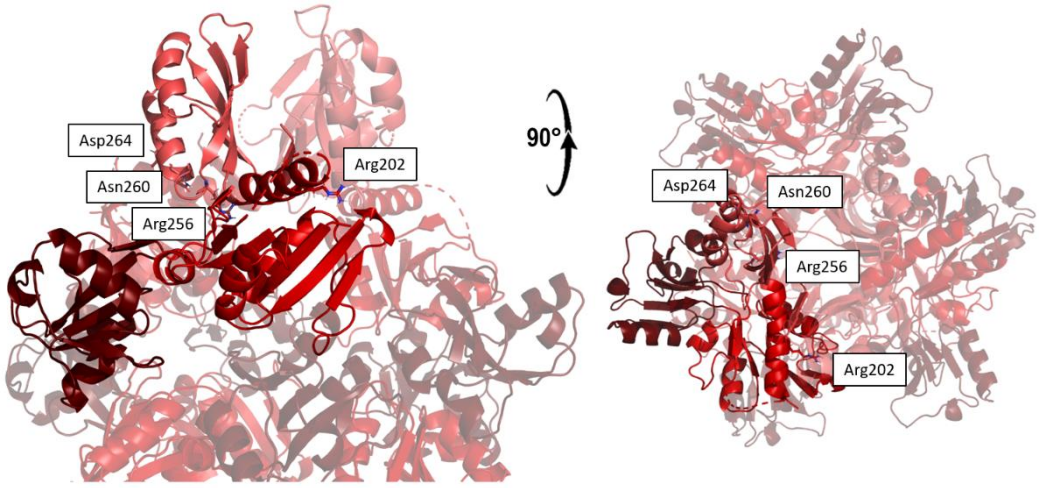

Figure 5.17. Residues of interest for *mtu*ATP-PRT as identified in *cje*ATP-PRT (Jiao *et al.*, 2019). Residues of interest are labelled on the hexameric structure of *mtu*ATP-PRT$_{WT}$ (PDB; 1NH7).

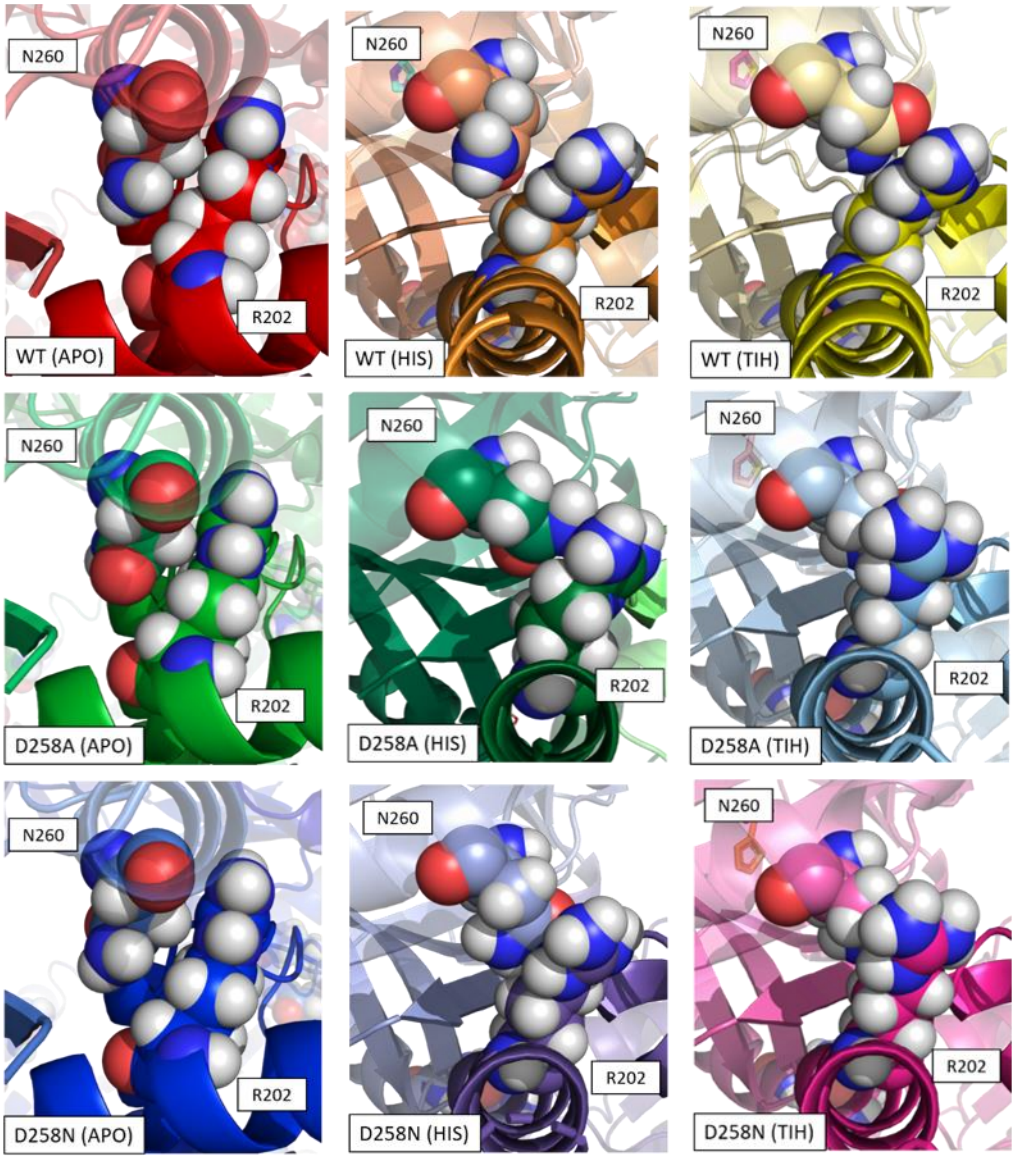

Figure 5.18. Interactions between residue Arg202 and Asn260 (spheres) described by Jiao *et al.* (2019) to be key residues involved in histidine inhibition in *C. jejuni*.

The interactions of each structure shown in **Figure 5.18** illustrate the slight changes to residue interactions on allosteric binding. In this interaction Arg202 on chain A interacts with Asn260 of chain B tightening the trimeric interface of domain III. The degree in which this distance between residues decreases is most notable in *mtu*ATP-PRT$_{D258A}$ and *mtu*ATP-PRT$_{D258N}$ TIH bound structures. But in *mtu*ATP-PRT variants there is no interaction switch of Arg256 (*cje*E217), which was the interaction observed in *cje*ATP-PRT under allosteric inhibition.

5.2.8 Interface residues potentially involved in allosteric activation

Phenix structure comparison was used to compare both previously determined structures of *mtu*ATP-PRT$_{WT}$ with allosteric ligands bound (PDB: 5LHU, 5LHT). This comparison identified a rotation of two residues unique to the *mtu*ATP-PRT$_{WT}$ activated structure. In the TIH bound conformation residue Glu18 rotates down, and Arg27 of Domain I rotates toward E18 (**Figure 5.19**) forming a salt bridge. These residues were then inspected in *mtu*ATP-PRT mutants in each allosteric state (**Figure 5.20**).

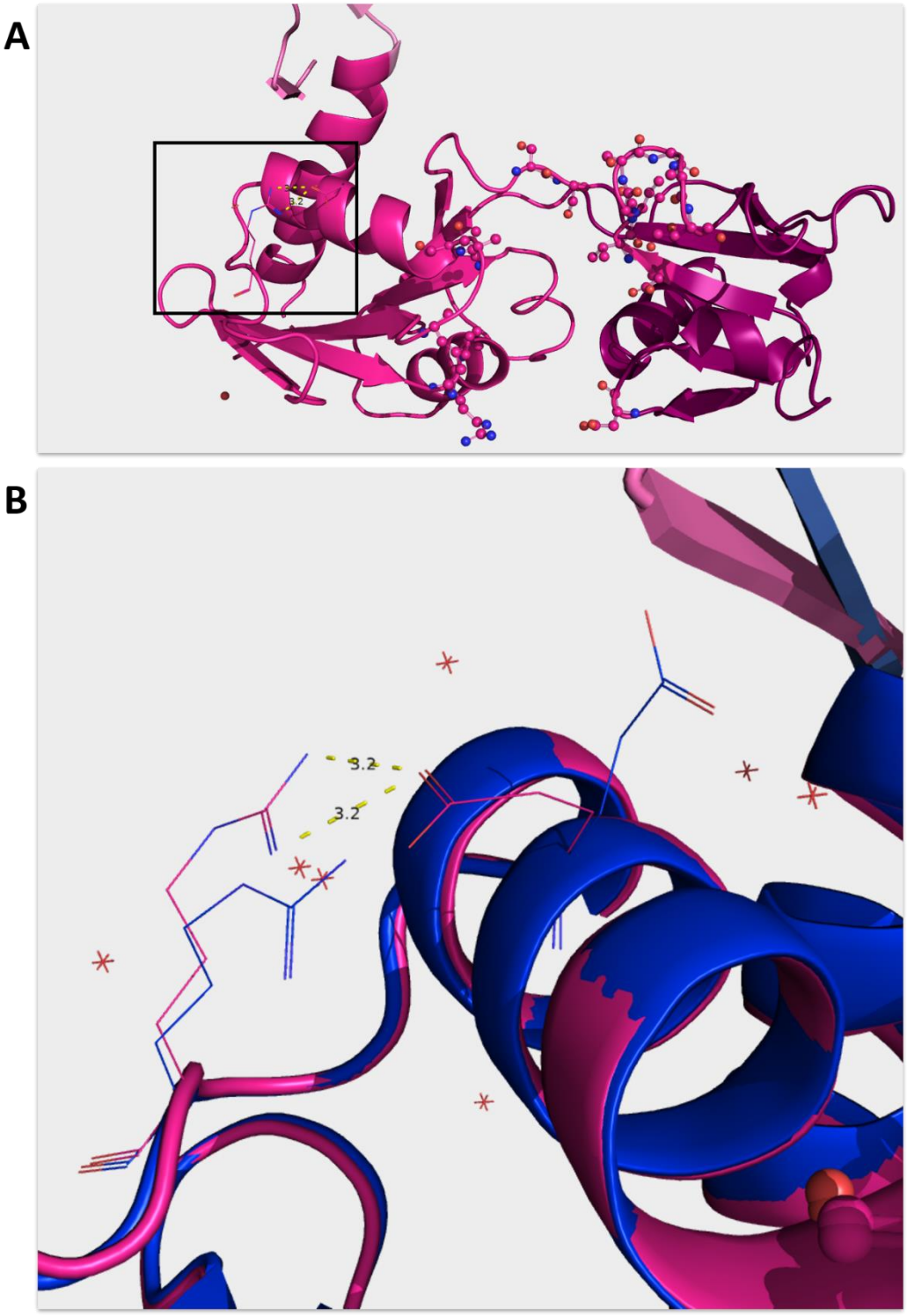

Figure 5.19. Rotating side chains Glu18 and Arg27 on allosteric activation by TIH. **A.** TIH bound *mtu*ATP-PRT$_{D258N}$ residue interactions. Active site residues are shown as ball and stick. **B.** Close comparative view of APO *mtu*ATP-PRT$_{D258N}$ (blue) and TIH *mtu*ATP-PRT$_{D258N}$ (magenta). In the TIH bound structure residues Glu18 and Arg27 interact at a distance of 3.2 Å.

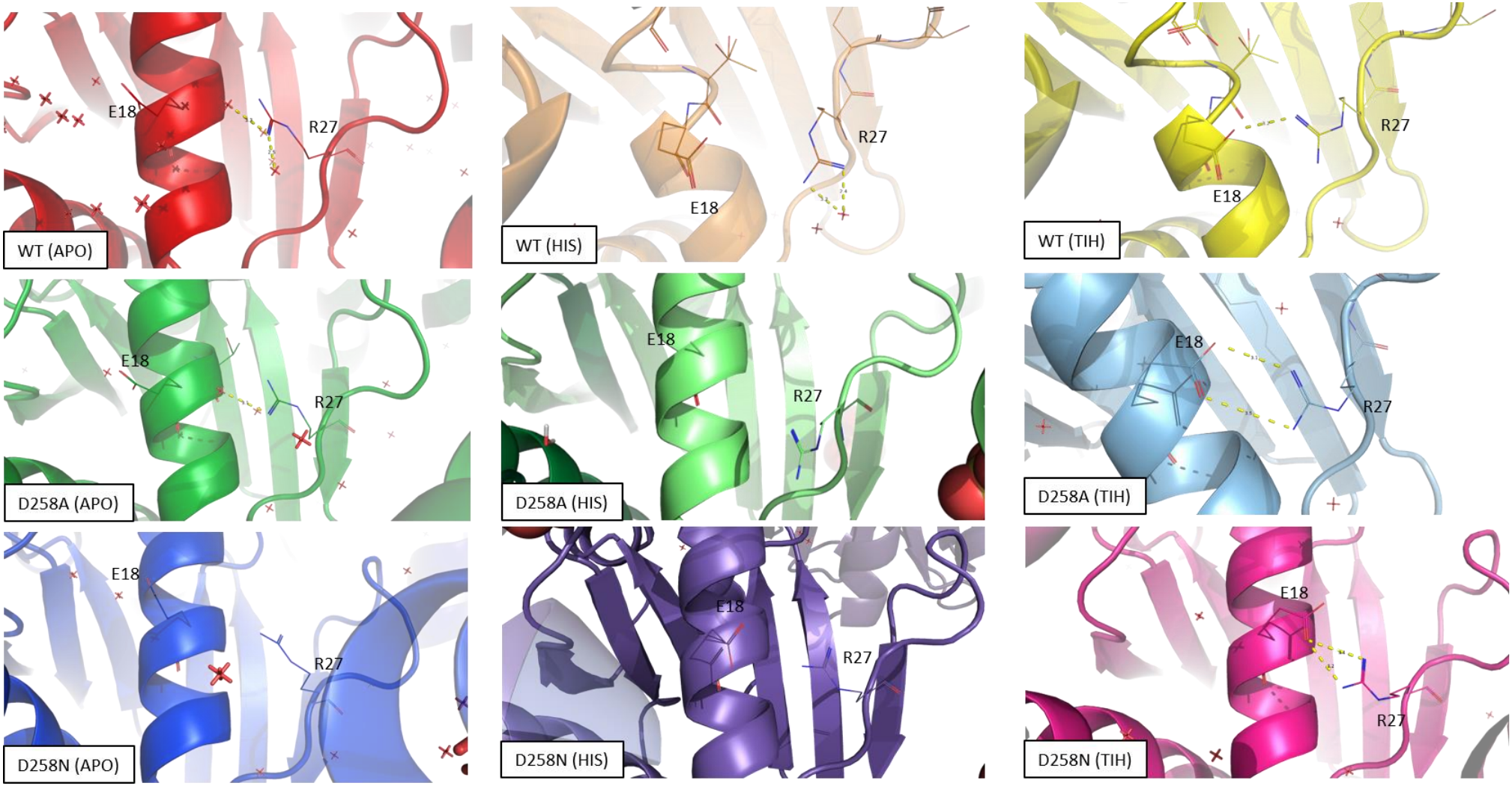
E18
R27
WT (APO)
E18
R27
WT (HIS)
E18
R27
WT (TIH)
E18
R27
D258A (APO)
E18
R27
D258A (HIS)
E18
R27
D258A (TIH)
E18
R27
D258N (APO)
E18
R27
D258N (HIS)
E18
R27
D258N (TIH)

Figure 5.20. Close structural analysis of residues Glu18 and Arg27 of *mtu*ATP-PRT$_{WT}$ (PDB: 1NH7, 5LHU, 5LHT), *mtu*ATP-PRT$_{D258A,}$ and *mtu*ATP-PRT$_{D258N.}$ Phenix Structure Compare identified a rotation of Glu18 and Arg27 upon binding of TIH. This rotation allows an interaction at 3.2 - 3.4 Å between residues Glu18 and Arg27 in the allosterically active state. This rotation was identified in each ATP-PRT variant in TIH co-crystallised structures and may facilitate increased stability in the active site for allosteric activation. Waters have been included in each figure to identify any water mediated bonds between residues Glu18 and Arg27.

Structure comparison of each *mtu*ATP-PRT variant co-crystallised with TIH identified the same rotation of both Glu18 and Arg27. This interaction is absent in structures with the inhibitory L-His ligand, and consistently present with the activating TIH ligand in all *mtu*ATP-PRT variants. In *mtu*ATP-PRT$_{D258N}$ (L-His insensitive mutant) co-crystallised with L-His there is some additional movement in Glu18 compared to the other structures co-crystallised with L-His, suggesting this residue movement could be key to ATP-PRT modulation. This movement of Glu18 orientates the residue toward Arg27, like in the TIH activated structures, however, Arg27 is oriented away from Glu18 in the same manner as seen in TIH activation, and thus the interaction cannot be formed.

5.2.9 Hydrogen bond analysis

Hydrogen bonds (H-bonds) were analysed in addition to analysis of structural changes between *mtu*ATP-PRT mutants to obtain a general overview of bonding across the structures. Hydrogens were added to the hexameric assembly of each *mtu*ATP-PRT structure and were cross-referenced against each allosteric state of each variant to extract unique hydrogen bond interactions in each crystal structure (**Figure 5.21**). That is, each structure (APO, L-His, and TIH) had all H-bonds identified. These H-bonds were categorised into three groups based on APT-PRT variant (i.e., WT APO, L-His, and TIH, *mtu*ATP-PRT$_{D258A}$ APO, L-His, and TIH, and *mtu*ATP-PRT$_{D258N}$ APO, L-His, and TIH). Each of the three groups were independently run through the H-bond pipeline, and all H-bonds unique to one allosteric state, compared to its variant counterpart were extracted resulting in nine datasets containing the unique H-bonds for each allosteric state for each ATP-PRT variant.

These unique bonds were coloured by chain interaction for clarity, as some bonds were between residues in the same chain (intra-chain), and others were to other chains (inter-chain) in the hexamer. Although the number of unique hydrogen bonds is large, the distribution and interaction within and between chains can be useful in the interpretation of signal transfer in large protein systems. The large number of data points provided by this H-bond analysis of affixed crystallographic structures provides a gross look at how hydrogen bonds change between allosteric states and upon mutation.

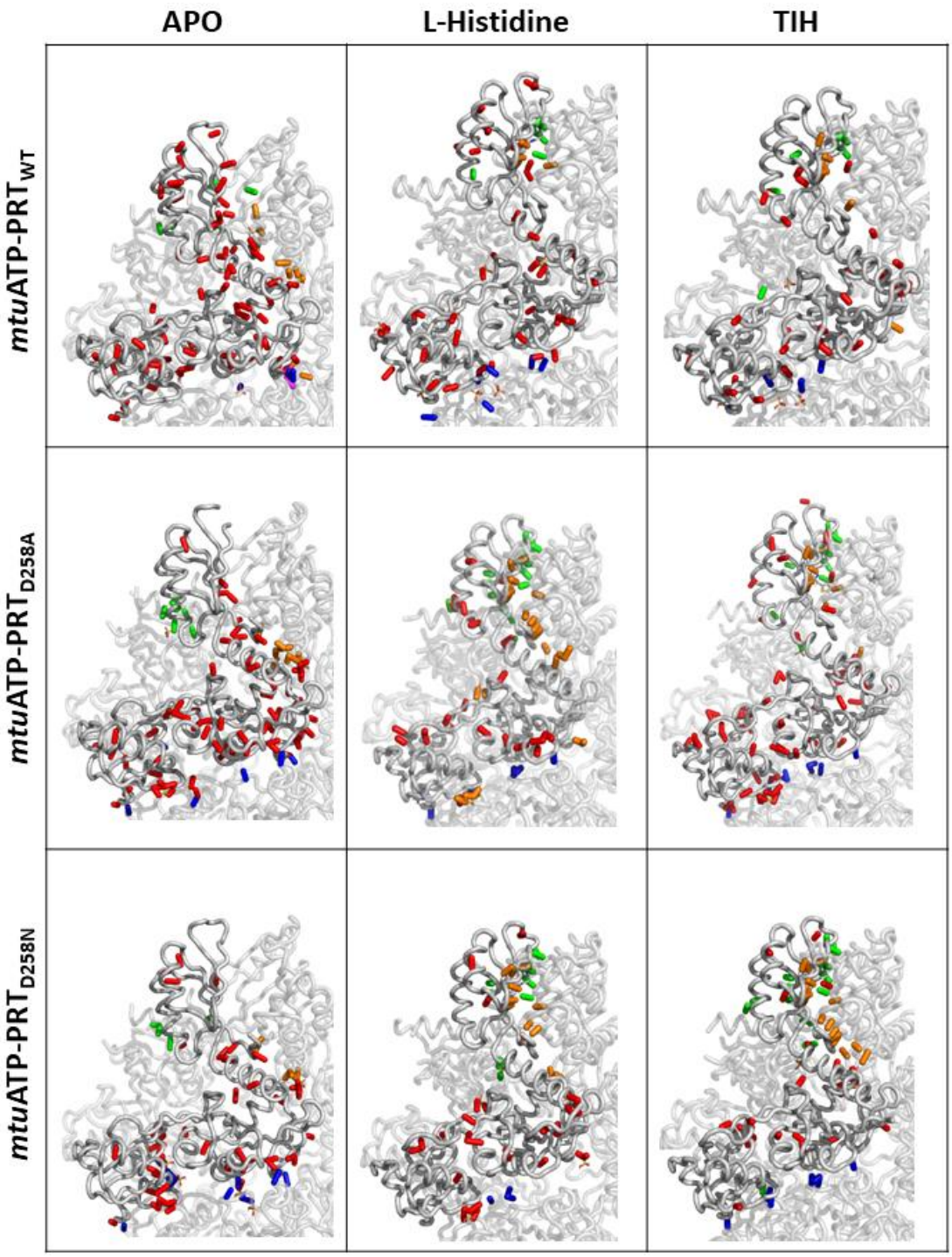

Figure 5.21. Hydrogen bond comparison of *mtu*ATP-PRT variants in each allosteric state. Hydrogen bond interactions are displayed as sticks, each is representative of a residue-residue interaction that is unique to the allosteric state of each respective ATP-PRT variant (ie; Bonds present in the APO bound structure of *mtu*ATP-PRT$_{WT}$ but not in the L-His or TIH bound structure of *mtu*ATP-PRT$_{WT}$). Bonds were coloured based on interaction with chain A; intra-chain interactions (red), chain B (orange), chain C (green) chain D (blue), and chain E (magenta). There were no unique interactions between chain A and F observed between allosteric states.

Intra-chain bonds are the most prevalent changed bonds across all the structures. *mtu*ATP-PRT$_{WT}$ APO showed a large number of these intra-chain bonds, along with a collection of bonds stabilising domains I and III of chain A to chain B (**Figure 5.21**). Three bonds interacting across the regulatory domain interact with chain C. Finally domain I of *mtu*ATP-PRT$_{WT}$ APO interacts with chain D and E through a small number of bonds near the final loop and helix of domain I. Overall, across all nine variants there are a greater number of unique interactions at the trimeric (allosteric) interface formed, compared to interactions at the dimeric interface, particularly under allosteric modulation. When bound with L-His and TIH, *mtu*ATP-PRT$_{WT}$ shows a similar number of intra-chain bonds in domains I and II, but fewer in domain III. Additionally there is a decrease in the number of interactions to chain B, and the bonds that remain are placed in domain III. An increase of interactions to chain C via both the α-helices and β-sheets accompany the remaining chain B interactions.

In the *mtu*ATP-PRT$_{D258A}$ structures, more notable changes to hydrogen bonding is observed upon ligand binding perhaps indicative of the exaggerated inhibitory and activatory effects in this mutant. The APO structure shows a number of unique bonds to chain C via the second α-helix of the regulatory domain. In the L-His bound form there are far fewer unique bonds present, indicating less change being propagated through the system, and reinforcing the anomalous nature of this structure. Compared to other *mtu*ATP-PRT$_{D258A}$ structures, that with histidine bound shows a substantial distribution of bonds interacting between chain A and B. Interactions to chain C are predominantly formed by interactions involving the β-sheets in domain III. With TIH bound this structure shows a number of unique hydrogen bonds interacting with the residues of chain A. Few interactions to chain B are seen compared to the histidine bound structure, as there is an increase of bonding to the C-terminus of chain C as the structure adopts the tensed state.

Finally in the *mtu*ATP-PRT$_{D258N}$ structures, the APO bound structure shows minimal unique hydrogen bonds, again very few are seen in the trimeric interface. There are minimal bonds interacting to chain B, and those that remain are involved with the α-helix of domain I. In the L-His bound structure there are substantially more interactions with chain B and chain C, predominantly in domain III. The TIH bound structure shows similar bonding patterns as L-His bound, indicating although

the H-bonds are unique between allosteric ligand bound states they are occurring in similar positions.

5.3 Discussion

The mutation position (258) lies at the end of a β-sheet and the beginning of a loop in the C-terminal regulatory domain of *mtu*ATP-PRT and mutations to residue sidechains were not expected to cause major changes to the protein structure, only to alter allosteric signal communication. As a result each mutant crystallised under similar conditions to *mtu*ATP-PRT$_{WT}$ (Cho *et al.*, 2003; Pisco *et al.*, 2017), space group, and each of the six solved structures maintained the same predicted protein structure as *mtu*ATP-PRT$_{WT}$ in either relaxed (APO) or tensed states (ligand-bound). All of these findings align with the minor change to protein sequence, and differences which may be found to explain kinetics changes due to mutation are expected to be found in the subtle differences of residue position and interaction within the structures. How these structural changes align to other approaches and the kinetic effects are discussed later in the thesis (Chapter 7).

5.3.1 Structure of Δ258 variants

The resolution for each structure were between 2.0 – 2.5 Å making them suitable for comparison without significant confounding effects from data quality differences. Each allosteric ligand was clearly identifiable in the trimeric interface by density, and increased density from the sulphur of TIH made it possible to orientate the thiophenyl ring with confidence.

Between the structures, there was variation in the cysteine bridge formation between Cys73 and Cys175 that bridges domains I and II. In *mtu*ATP-PRT$_{D258N}$ (APO) there was no evidence of a disulphide bridge, however in all other structures the density supports this bond. All three TIH bound structures showed evidence of alternate conformations for oxidised and reduced interactions. The variation of bridge presence or absence is unlikely to be caused by exposure to different concentrations of reducing agents (present in TEV digest section 3.3.1) during purification. Each protein variant was batch purified, and thus differences in Cys73 – Cys175 interactions are not the result of purification conditions. In terms of

function this disulphide bond between domain I and II of the active site would provide rigidity surrounding the catalytic domain, and potentially stabilise the protein. This bond could result in reduced catalysis, if increased rigidity restricted the protein, particularly as it is located through the catalytic domain. Therefore, both TIH bound states showing alternate conformations of Cys73 and Cys175 could contribute to the increased catalysis seen in TIH modulated protein *via* releasing the restriction through the catalytic core. The alternating conformations suggests activation changes the frequency in which this bond is formed. Furthermore, this interaction was absent in APO *mtu*ATP-PRT$_{D258N}$ (**Figure 5.10 A**) which showed increased catalysis without allosteric modulator (**Figure 4.12 C**). If this bridge contributed to enzyme rigidity, and thus reduced catalysis, the addition of reducing agents during assay may be able to undo this interaction and increase catalysis. However, *in vivo* the reducing environment of the cell may reduce these disulphides, suggesting the presence and absence of these is likely a function of the *ex vivo* crystallographic conditions (Carmel-Harel & Storz, 2000)

5.3.2 Structural comparisons

Long form ATP-PRT adapts a tensed state under L-His and TIH modulation seen in PDB structures 5LHU and 5LHT (Pisco *et al.*, 2017). Global RMSD analysis of ATP-PRT variants showed low RMSD (and thus high structural similarity) between variants and their respective allosteric states (**Figure 5.14**). When each domain was extracted and overlaid, it was identified that domain II moved the least between all allosteric states. Domain II encompasses approximately half of the active site residues, however the movements between tensed and relaxed states are minimal compared to domains I and III which move dramatically on multiple planes in the tensed state. On all RMSD graphs (**Figure 5.14 A**, **B**, **C**, and **D**) it is evident how similar each L-His and TIH bound structures are to one another despite their diametrically opposed effects on catalysis. This is not unexpected in the global RMSD, but it was surprising to see increased variation in domain II movements between L-His and TIH bound states of Δ285 variants considering domain II is not heavily involved in the movement between tensed and relaxed states of the enzyme. However, the severity of this movement is not as marked as differences between domain I or III movements, as reflected by the lowered variation in RMSD within all domain II residues **Figure 5.14**. Furthermore on closer inspection, the

differences between the structure lay in loop regions of the domain. The subtle RMSD change between the variants suggests kinetic and allosteric changes are attuned differences that may not be discovered with coarse comparison of static crystal structures, and instead require close comparison of individual residue movements.

Pisco *et al.* (2017) proposed a tightening between residues Pro255 and Asp258 was a mechanistic switch of L-His inhibition. Whilst we saw interactions of these residues certainly increased under allosteric modulation in wild type (**Figure 5.16**), these interactions were altered by mutations to position 258. Naturally altering the charge of residue 258 is going to alter the interactions to nearby residues, however, in *mtu*ATP-PRT$_{D258A}$, L-His and TIH both elicit exaggerated allosteric effects, even with reduced interactions to residue Pro255. In contrast, in *mtu*ATP-PRT$_{D258N}$, allosteric inhibition by L-His is lost, along with interaction to Pro255, but TIH activation remains, even without this interaction in place. This suggests although there is a recognised interaction between these residues it is unlikely to be a driving mechanism in allosteric communication. Indeed, these static interactions may not provide the necessary insights if the mutations are, in fact, eliciting changes to the dynamics of the system.

*cje*ATP-PRT residues involved in L-His inhibition have been previously identified by Jiao *et al.* (2019). These residue interactions are predicted to be involved in the stabilisation of the regulatory domain. These *cje* residues were cross checked in *mtu*ATP-PRT variant structures. The *mtu* equivalent residues did not show the same effect in wildtype or allosteric states. Arg202 and Asn260 were predicted to dissociate during allosteric inhibition by L-His, and instead Arg202 would establish contact with Arg256, thus destabilising the regulatory domain. Instead this initial contact remained across all allosteric states and was more pronounced in both *mtu*ATP-PRT$_{D258A}$ and *mtu*ATP-PRT$_{D258N.}$ Although structurally similar, *cje* and *mtu* share merely 30 % sequence identity. In *cje* the charge of these residues may contribute to their association, however, the equivalent residues in *mtu* all held a positive or no charge. The low sequence similarity between *cje* and *mtu* indicates a long period of evolution between the two species (Chapter 2), which may also drive alternative residue interactions involved in allosteric control.

Figure 5.19 and **Figure 5.20** show a rotamer identified in all TIH activated structures where Glu18 and Arg27 appear to move to positions where they may interact. This interaction would provide stability between the loop helix of domain I, on the edge of the active site. This interaction may provide stability to the active site thus assisting in increased catalysis. However, there is no evidence as to what triggers this interaction to occur from TIH binding in domain III. Regardless of communication, this stabilisation offers preorganization to the catalytic site of the enzyme.

Hydrogen bond analysis of all *mtu*ATP-PRT structures provides insights into residue interactions unique to each allosteric state and each mutant. This comparative analysis (**Figure 5.21**) identified loss and gain of hydrogen bonds particularly between tensed and relaxed states of *mtu*ATP-PRT. In each variant there was an increase of H-bonds forming inter-chain interactions (chain A to chain B, and chain A to chain C) in domain III, tightening the trimeric interface (green sticks). Comparing WT L-His and *mtu*ATP-PRT$_{D258A}$ L-His (**Figure 5.21**) there is significant reduction to the number of inter-chain bonds formed, particularly in domain I and II. We see a reduction of these inter-chain H-bonds in L-His insensitive *mtu*ATP-PRT$_{D258N}$ compared to L-His inhibited WT. Whilst the C-terminal regulatory domain H-bonds appear similar, the significant reduction of bonds through the domains I and II of the active site may highlight suggest the enzyme is not adopting the typical **T** state expected from L-His inhibition. Both allosteric modulators in all variants propagated increased interactions in the allosteric domain, which relate to the movement into the tensed state. *mtu*ATP-PRT$_{D258A}$ with L-His bound showed a reduction in bonds, additionally, this variant had the lowest catalytic rates (**Figure 4.12**) and was super-sensitive to L-His inhibition at low temperatures. B-factor results (**Figure 5.12**) showed *mtu*ATP-PRT$_{D258A}$ structured had the highest average B-factor. The only exception being residues 230 - 250 of domain III residues where the average B-factor was reduced relative to previous B-factors in *mtu*ATP-PRT$_{D258A}$ structures.

High B-factors represent flexible ordered regions of protein, considering the near identical protein sequence we would expect each B-factor to track similarly across the protein chain. For each ATP-PRT variant with and without allosteric modulator we see their B-factors align in many regions (with the exception of very low

resolution *mtu*ATP-PRT$_{WT}$ APO). However, these diverge further in the final 70 residues (that comprise domain III) suggesting the site mutations at 258 may have an effect on the flexibility and order in the regulatory domain.

Through structural analysis alone we were unable to determine the full communication pathways *mtu*ATP-PRT is using for allosteric regulation across the 40 Å distance of the structure. This is unsurprising given the dynamic information that is lost in crystallographic structures, however key changes along the pathways have been identified. There are multiple residues that move dependent on allosteric state, and these may contribute to altered catalysis. These include Glu18 – Arg27 interaction upon TIH binding, disulphide bridge formation between Cys73 – Cys175 which demonstrates reduced occupancy upon TIH binding. Allosteric regulation is a highly dynamic process that may not be visible in a snapshot crystallographic state, but the differences discovered here provide insight into the communication chain and provide a guide for further MD work to investigate the dynamic contribution to allostery in ATP-PRT.

5.4 Conclusions

The protein structures of two *mtu*ATP-PRT mutants in three allosteric states have been solved at good resolution. In conjunction with existing wildtype structures, these provide a suite of crystallographic structures for use in the investigation of allosteric change in ATP-PRT. The four structures with allosteric ligands bound support previous evidence that both allosteric activation and inhibition elicit an identical tensed form in *mtu*ATP-PRT, however this contradicts the notion that tension of an enzyme drives catalytic rate. Although these findings could not elucidate a chain of crucial residues involved in allosteric communication, we identified a number of key residues varying between active and inhibited states. Looking for allosteric inhibition changes that may be associated with histidine insensitivity in *mtu*ATP-PRT$_{D258N}$ uncovered a potential interaction involved in allosteric activation by TIH. These residues Glu18 and Arg27 may serve to stabilise the active site during catalysis and could be further investigated by additional site directed mutagenesis altering these residues to either force or inhibit interactions. Hydrogen bond analysis identified large numbers of unique hydrogen bonds formed

between allosteric states that change upon ligand binding. The patterning of these hydrogen bonds is representative of domain shifts and the tensing of the overall structure. This chapter suggested allostery in *mtu*ATP-PRT is a highly dynamic process that cannot be determined using crystallographic evidence alone, however, has produced a number of crystal structures for additional molecular dynamics investigations to add to our knowledge of *mtu*ATP-PRT allosteric regulation.

Chapter 6

Preliminary molecular dynamics analysis of *mtu*ATP-PRT variants in allosteric conformational states

Computational biology *via* molecular dynamic (MD) simulations of proteins provides a novel method of examining subtle allosteric changes in complex enzyme systems. MD simulations have been used successfully in the investigation of *cje*ATP-PRT allosteric inhibition by L-His (Jiao *et al.*, 2019), however, have not been applied to *mtu*ATP-PRT, nor in the context of TIH activation. Crystal structures of *mtu*ATP-PRT show L-His and TIH stimulate the same tensed crystallographic state of the enzyme. Crystallography only provides a snapshot of a single conformation thus, crystallographic data can only provide clues to differences between position 258 mutants, and the allosteric states across the range of systems examined. These crystallographic data cannot represent the dynamic changes that occur during an enzyme reaction. MD samples multiple conformations and simulates protein dynamics that may be central in allosteric modulation. The work described in this section is a collaboration between myself and Dr Wanting Jiao at Victoria University of Wellington. Dr Jiao took the structures described in Chapter 5 and set up the MD simulations. The preliminary results were analysed jointly and are discussed below.

6.1 Methods

APO, L-His, and TIH bound crystal structures *mtu*ATP-PRT$_{D258N}$ from Chapter 5 and *mtu*ATP-PRT$_{WT}$ PDB; 1NH7, 5LHU, and 5LHT were prepared for MD simulations to investigate subtle dynamic changes to *mtu*ATP-PRT variants in different allosteric states. Each structure had backbones filled (**Figure 5.7**) and was assembled into their biological hexamer using symmetry mates operations in PyMOL. These structures were shared, and preliminary MD data was kindly collected by Dr Jiao (Ferrier Research Institute, Victoria University of Wellington)

using Desmond (Bowers *et al.*, 2006). Protein backbone RMSD, and hydrogen-bonding patterns were analysed, with MD simulations collected at 310 K.

6.2 Results

Molecular dynamics simulations were run in triplicate for 550 – 1594 ns periods described in **Table 6.1** with between 1.2 – 2.5 µs total equilibrated time periods collected for each allosteric state of *mtu*ATP-PRT$_{WT}$ and *mtu*ATP-PRT$_{D258N}$ (*mtu*ATP-PRT$_{D258A}$ was not collected for preliminary MD due to time constraints). Differences in run time were due to the varying lengths of time individual systems took to reach equilibrium (a stable RMSD; **Figure 6.1**). Δ258N with L-His was collected for significantly longer than any other state due to this equilibration time.

Table 6.1. MD set up run times. MD simulations were run on *mtu*ATP-PRT$_{WT}$ and *mtu*ATP-PRT$_{D258N}$ in APO, L-His, and TIH bound states. Each structure was simulated in triplicate for varying lengths (as listed) based on gathering sufficient equilibrated data. Between 1.2 and 2.5 µs of equilibrated data was collected for each system.

	WT APO	**WT+HIS**	**WT+TIH**	**D258N APO**	**D258N+HIS**	**D258N+TIH**
1	649 ns	550 ns	600 ns	606 ns	1594 ns	597 ns
2	600 ns	600 ns	600 ns	600 ns	1500 ns	500 ns
3	700 ns	1300 ns	600 ns	598 ns	1500 ns	900 ns
Total equilibrated time	1.8 µs	1.7 µs	1.4 µs	1.3 µs	2.5 µs	1.2 µs

Following the equilibrium of the simulation the RMSD values of backbones were annotated for each system (**Figure 6.1**). The RMSD for each was collected between the marked sections of **Figure 6.1** and plotted in **Figure 6.2**.

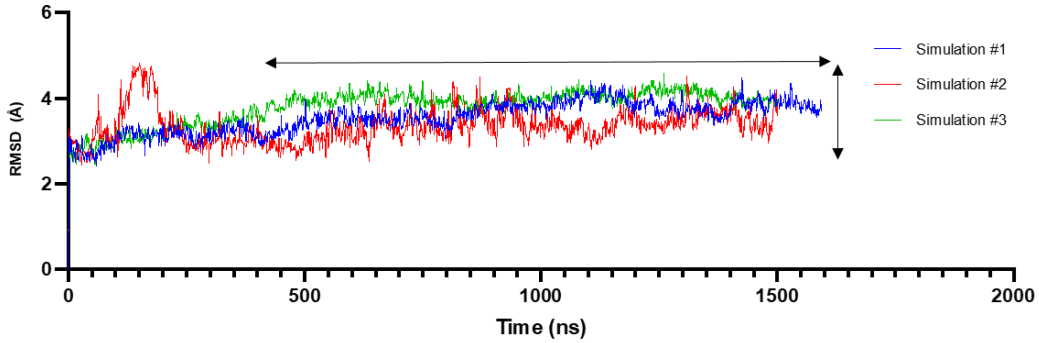

Figure 6.1. Protein backbone RMSD values during MD simulations. Horizontal arrows indicate equilibrated, sampled data used to generate **Figure 6.2**. Vertical arrow represents the variation to RMSD.

Figure 6.2 illustrates the variation of the average RMSD in the backbone of each protein collected from MD simulations (**Table 6.1**). The error bars represent the variation in RMSD across the simulation, and thus reflect the fluctuation of movement and flexibility of the enzyme during simulation. The average RMSD represents the overall flexibility of the enzyme, and error bars represent how this RMSD changed over the course of the simulation.

In both APO structures, *mtu*ATP-PRT$_{WT}$ and *mtu*ATP-PRT$_{D258A}$ show similar mean RMSD and error, indicating similar levels of flexibility in this state. We begin to observe changes to protein flexibility in the binding of L-His, where *mtu*ATP-PRT$_{WT}$ displays lowered variation to backbone RMSD, and thus reduced flexibility. This variation to RMSD is greater in the *mtu*ATP-PRT$_{D258N}$ mutant, which kinetically (**Figure 4.12**) shows resistance to L-His inhibition. There is a stark difference upon TIH binding observed between *mtu*ATP-PRT$_{WT}$ and *mtu*ATP-PRT$_{D258N}$. One WT simulation shows a much higher RMSD with greater error compared to any other *mtu*ATP-PRT$_{WT}$ (TIH) or *mtu*ATP-PRT$_{D258N}$ (TIH) runs. However, overall, we see more variation between each TIH bound enzyme, with alterations to mean RMSD and flexibility surrounding these.

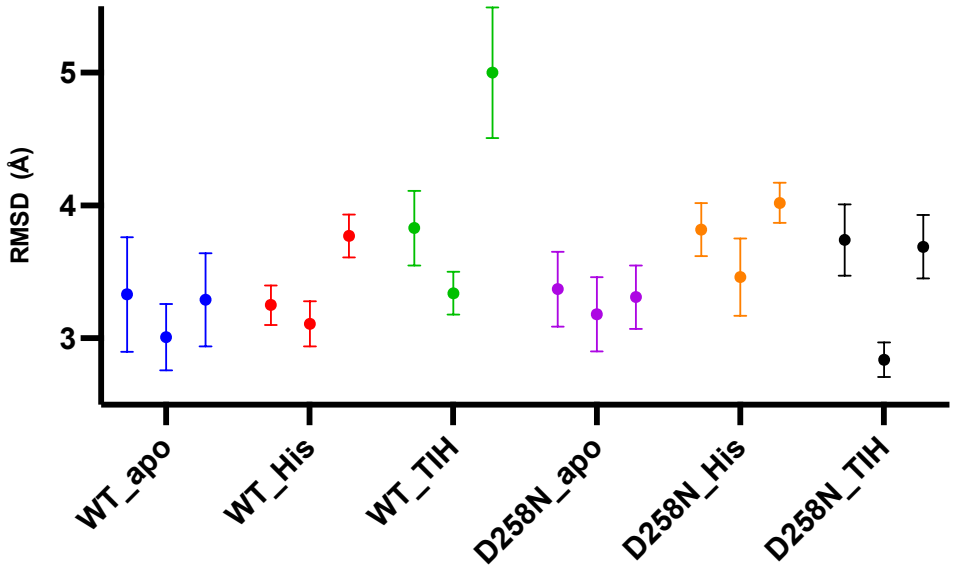

Figure 6.2. Average RMSD of protein backbones across three simulations of each protein in each allosteric state. Error bars indicate variation in RMSD, and thus the flexibility of the backbone during simulation.

MD simulations were also able to provide data on hydrogen bond changes across the three allosteric states in both *mtu*ATP-PRT variants. These can be examined to indicate what bonds are more sampled in different allosteric states. In Tables 1.2, 1.3 and 1.4, bonds broken in greater than 10 % of occupancies are coloured in red, and bonds formed in greater than 10 % of occupancies are coloured in green. Greater or lesser than 10 % indicates greater fluctuation of these bonds. Full H-bond tables can be found in Appendix 5 (**Table A.5.1, A.5.2, A.5.3**).

In **Table 6.2** we can see in *mtu*ATP-PRT$_{WT}$ the shared formation of four hydrogen bonds under allosteric modulation that are not present in APO protein. Glu61 – Leu146 and Ile239 – Pro241 both show the greatest occupancy of bond formation, and thus occur a greater propensity of the time. There is also the shared loss of four bonds under allosteric modulation that are present in APO protein, these have lower occupancy compared to formed bonds. We also observe bond changes between L-His and TIH bound states; with L-His bound enzyme forming two unique bonds not shared with TIH bound states; Arg27 – Glu18 and Asp278 – Asp278 (inter-chain). It is noteworthy that the Arg27-Glu18 interaction was previously identified in the static crystallographic structures.

In **Table 6.3** *mtu*ATP-PRT$_{D258N}$ we see a lesser number of hydrogen bond changes compared to *mtu*ATP-PRT$_{WT}$. The three bonds with increased frequency of formation are shared with *mtu*ATP-PRT$_{WT}$, suggesting commonality in allosteric modulated H-bond formation between variants. There is lowered formation of four bonds, with Arg280 – Asp278 being the most significant bond change which does not present in allosteric modulated states. Of these four reduced bonds, two (Ala276 – Phe281 and Glu198 – Asp264) are shared with *mtu*ATP-PRT$_{WT}$. In the L-His insensitive *mtu*ATP-PRT$_{D258N}$ there is only one significant hydrogen bond change observed in L-His state that is not present in TIH bound state also. Arg280 – Ala276 had a greater formation frequency in L-His bound states (44.8 % versus 31.7 %).

Table 6.4 shows cross checks between the unique bonds that form in L-His bound states of both *mtu*ATP-PRT$_{WT}$ and Δ258N. L-His sensitive WT showed higher propensity for the formation of three bonds. However, in Δ258N there was only an increased formation of one interaction; Arg280 – Ala276 that was more frequently observed in the L-His bound state. These reduced and formed bond formations in the L-His insensitive Δ258N variant may contribute to allosteric insensitivity. A reduction in H-bonds in L-His insensitive Δ258N was also observed in H-bond analysis of static crystal structures (**Figure 5.21**).

Table 6.2. Hydrogen bond analysis of *mtu*ATP-PRT$_{WT}$ from MD simulations averaged from three runs. Occupancy changes greater than 10% are marked in green (bonds present in a greater propensity of time) and red (bonds absent in a greater propensity of time). Full hydrogen bond analysis tables with changes less than 10 % can be found in Appendix 5 (**Table A.5.1**).

Donor residue	Acceptor residue	WT (APO)	WT (L-His)	WT (TIH)	L-His - APO	TIH - APO	L-His - TIH
GLU61-Main	LEU146-Main	0.0%	46.3%	46.3%	46.3%	46.3%	0.1%
ILE239-Main	PRO241-Main	0.3%	41.3%	37.1%	41.0%	36.9%	4.2%
ARG280-Main	ALA276-Main	47.7%	24.5%	22.1%	-23.2%	-25.6%	2.4%
ARG251-Side	ASP216-Side	10.6%	22.9%	23.1%	12.2%	12.4%	-0.2%
ARG283-Main	ILE274-Main	33.0%	22.8%	20.6%	-10.2%	-12.3%	2.2%
ARG27-Side	GLU18-Side	0.0%	17.3%	5.0%	17.3%	5.0%	12.3%
SER236-Side	ASP244-Side	1.1%	17.1%	17.2%	16.0%	16.1%	-0.1%
ARG283-Side	ASP264-Side	24.7%	12.7%	14.8%	-12.0%	-10.0%	-2.1%
ALA276-Main	PHE281-Main	26.3%	12.6%	12.6%	-13.7%	-13.7%	0.0%
ASP278-Main	ASP278-Main	8.5%	11.5%	0.1%	3.0%	-8.4%	11.4%
GLN198-Side	ASP264-Side	24.8%	1.2%	3.3%	-23.6%	-21.4%	-2.1%

Table 6.3. Hydrogen bond analysis of *mtu*ATP-PRT$_{D258N}$ from MD simulations averaged from three runs. Occupancy changes greater than 10 % are marked in green (bonds present in a greater propensity of time) and red (bonds absent in a greater propensity of time). Full hydrogen bond analysis tables with changes less than 10 % can be found in Appendix 5 (**Table A.5.2**).

Donor residue	Acceptor residue	Δ258N (APO)	Δ258N (L-His)	Δ258N (TIH)	L-His - APO	TIH - APO	L-His - TIH
GLU61-Main	LEU146-Main	0.0%	50.4%	50.6%	50.4%	50.6%	-0.2%
ILE239-Main	PRO241-Main	0.8%	17.7%	22.4%	16.9%	21.6%	-4.7%
ARG280-Main	ALA276-Main	33.7%	44.8%	31.7%	11.1%	-2.0%	13.1%
ALA276-Main	PHE281-Main	44.8%	22.5%	14.1%	-22.3%	-30.8%	8.5%
ASP278-Main	ASP278-Main	0.0%	12.2%	12.1%	12.2%	12.1%	0.1%
ARG256-Side	GLN210-Main	21.4%	8.4%	7.7%	-12.9%	-13.7%	0.8%
ARG280-Side	ASP278-Side	39.8%	5.8%	7.7%	-34.0%	-32.1%	-1.9%
GLN198-Side	ASP264-Side	22.1%	7.6%	5.4%	-14.5%	-16.7%	2.2%

Table 6.4. Comparison of *mtu*ATP-PRT$_{WT}$ and *mtu*ATP-PRT$_{D258N}$ L-His bound structures from MD simulation. Occupancy changes greater than 10% are marked in green (bonds present in a greater propensity of time) and red (bonds absent in a greater propensity of time). Full hydrogen bond analysis tables can be found in Appendix 5 (**Table A.5.3**).

Donor residue	Acceptor residue	WT (L-His)	Δ258N (L-His)	Δ258N (L-His) - WT (L-His)
ILE239-Main	PRO241-Main	41.3%	17.7%	-23.6%
ARG280-Main	ALA276-Main	24.5%	44.8%	20.4%
ARG27-Side	GLU18-Side	17.3%	3.3%	-14.0%
SER236-Side	ASP244-Side	17.1%	6.9%	-10.2%

Bond formation and loss from **Table 6.2**, and **Table 6.4** is visualised in **Figure 6.3**, **Figure 6.4** and **Figure 6.5**. **Figure 6.3 A** and **B** illustrate the bonds formed and broken in APO state of *mtu*ATP-PRT$_{WT}$ that alter from the allosteric ligand bound states (**C**, **D**, **E** and **F**). **Figure 6.3 A** shows a large number of bonds formed around domain I and domain III, and an absence of domain III to domain III interactions, like those seen in the allosteric ligand bound interactions. With allosteric ligand bound (**C** and **E**), there is a large number of bonds formed toward the top of domain III, consistent with the established tightened enzyme movement. Additionally, there is a section of bonds formed at the dimeric interface between chain A and D (coloured cartoon) with both L-His and TIH binding. **Figure 6.3 C** differs from **E** by the addition of Glu18 and Arg27 interaction.

Figure 6.4 shows far fewer bonds formed with significant propensities between allosteric states. The bonds that remained formed, and absent were very similar to WT. There was the introduction of bond formation (Ala276 and Phe281 in domain III) in **Figure 6.4 C** that appears unique compared to other variants and allosteric states. There is reduced bond formation in the C-terminus in figure **E** (TIH bound) compared to figure **C** (L-His) particularly in the region between domain I and domain III of the next chain.

Kinetic data of *mtu*ATP-PRT$_{WT}$ and *mtu*ATP-PRT$_{D258N}$ showed different responses to L-His (**Figure 4.12**). No major structural change was identified in analysis of the enzyme structure (**Figure 5.14**). **Figure 6.5** illustrates the difference in bonds between these two structures upon L-His binding as identified through MD simulations. **Figure 6.5 A** (*mtu*ATP-PRT$_{WT}$) shows an increased number of bonds formed upon L-His binding compared to **Figure 6.5 B** (*mtu*ATP-PRT$_{D258N}$). Panel **B** contains fewer unique bonds compared to Panel **A**, with increased propensity for only Arg280 and Ala276 bond interactions. This indicates bond formations around the regulatory domain (seen at a greater degree in panel **A**) are characteristic of T-state ATP-PRT are reduced in the L-His insensitive Δ258N variant, consistent with the observed broken His communication pathways in this variant.

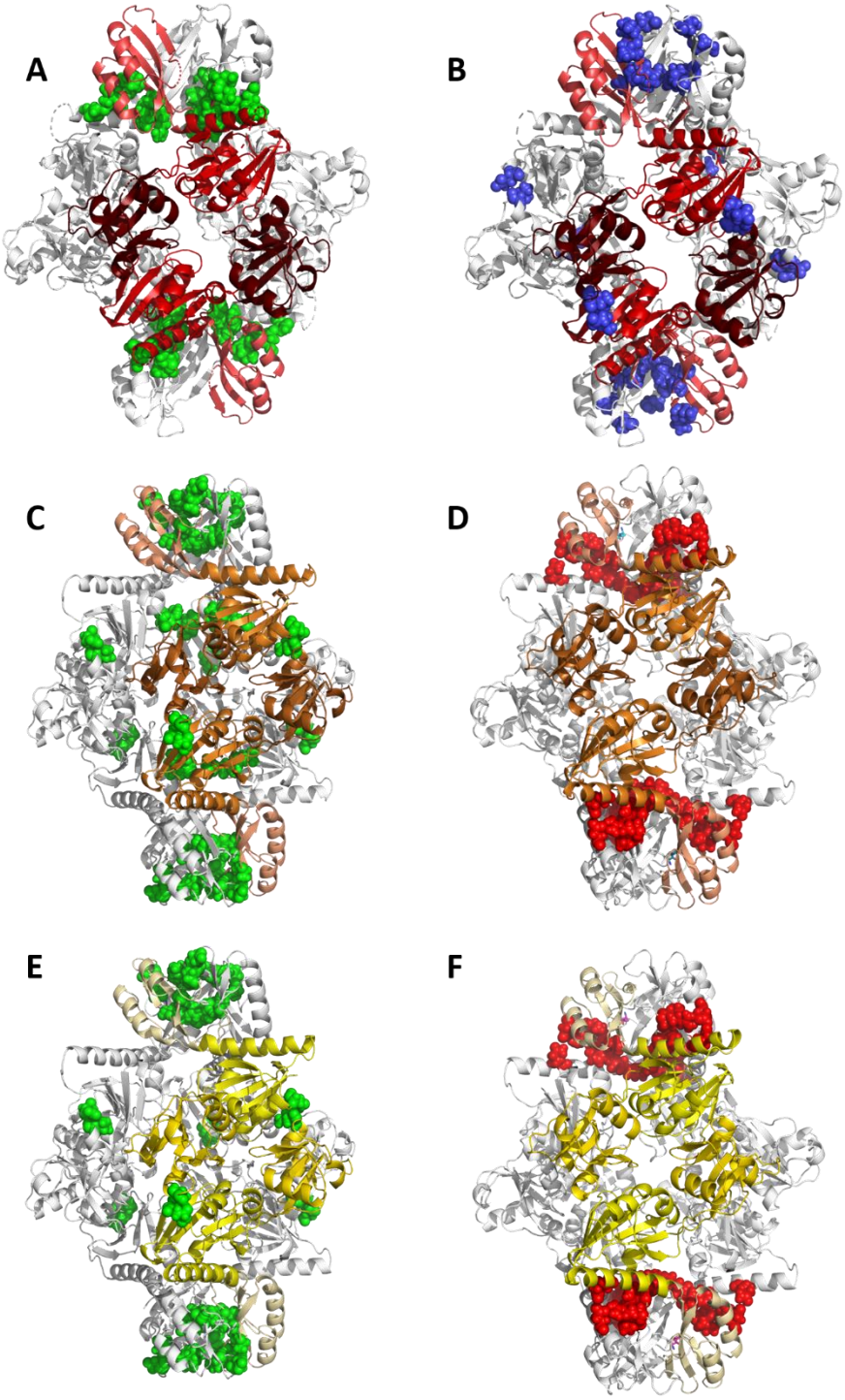

Figure 6.3. Illustrated H-Bonds formed (green spheres) and broken (blue or red spheres) in *mtu*ATP-PRT$_{WT}$ between allosteric states. Bonds formed or broken are defined as interactions that occur in greater or lesser propensity (10% frequency change compared to other allosteric states) as identified in MD simulations (**Table 6.2**). **A.** *mtu*ATP-PRT$_{WT}$ APO bonds formed in APO state compared to *mtu*ATP-PRT$_{WT}$ L-His. **B.** *mtu*ATP-PRT$_{WT}$ APO bonds broken in APO state compared to *mtu*ATP-PRT$_{WT}$ L-His. **C.** *mtu*ATP-PRT$_{WT}$ L-His bonds formed compared to *mtu*ATP-PRT$_{WT}$ APO state. **D.** *mtu*ATP-PRT$_{WT}$ L-His bonds broken compared to *mtu*ATP-PRT$_{WT}$ APO state. **E.** *mtu*ATP-PRT$_{WT}$ TIH bonds formed compared to *mtu*ATP-PRT$_{WT}$ APO state. **F.** *mtu*ATP-PRT$_{WT}$ TIH bonds broken compared to *mtu*ATP-PRT$_{WT}$ APO state.

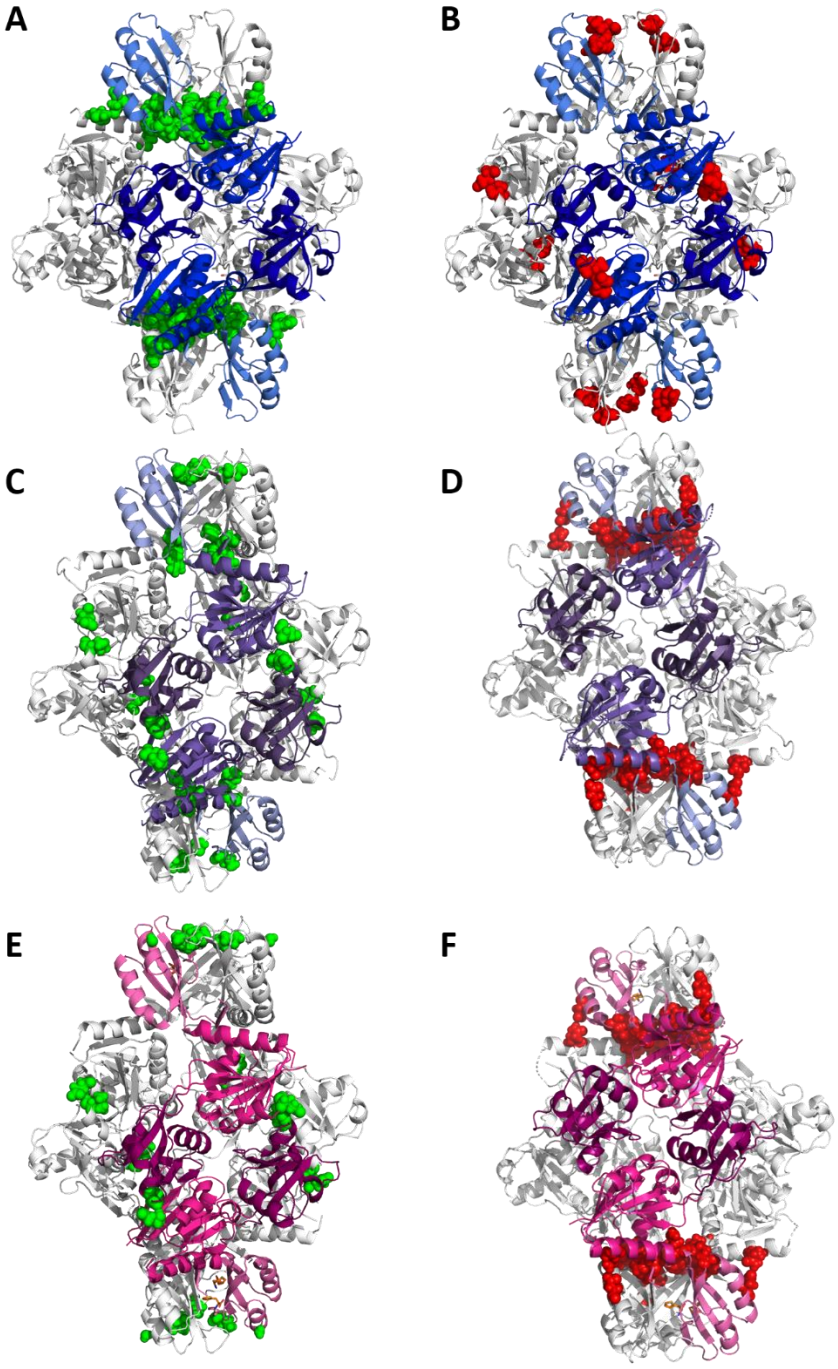

Figure 6.4. Illustrated H-Bonds formed (green spheres) and broken (red spheres) in *mtu*ATP-PRT$_{D258N}$ between allosteric states. Bonds formed or broken are defined as interactions that occur in greater or lesser propensity (10% frequency change compared to other allosteric states) as identified in MD simulations (**Table 6.3**). **A.** *mtu*ATP-PRT$_{D258N}$ APO bonds formed in APO state compared to *mtu*ATP-PRT$_{D258N}$ L-His. **B.** *mtu*ATP-PRT$_{D258N}$ APO bonds broken in APO state compared to *mtu*ATP-PRT$_{D258N}$ L-His. **C.** *mtu*ATP-PRT$_{D258N}$ L-His bonds formed compared to *mtu*ATP-PRT$_{D258N}$ APO state. **D.** *mtu*ATP-PRT$_{D258N}$ L-His bonds broken compared to *mtu*ATP-PRT$_{D258N}$ APO state. **E.** *mtu*ATP-PRT$_{D258N}$ TIH bonds formed compared to *mtu*ATP-PRT$_{D258N}$ APO state. **F.** *mtu*ATP-PRT$_{D258N}$ TIH bonds broken compared to *mtu*ATP-PRT$_{D258N}$ APO state.

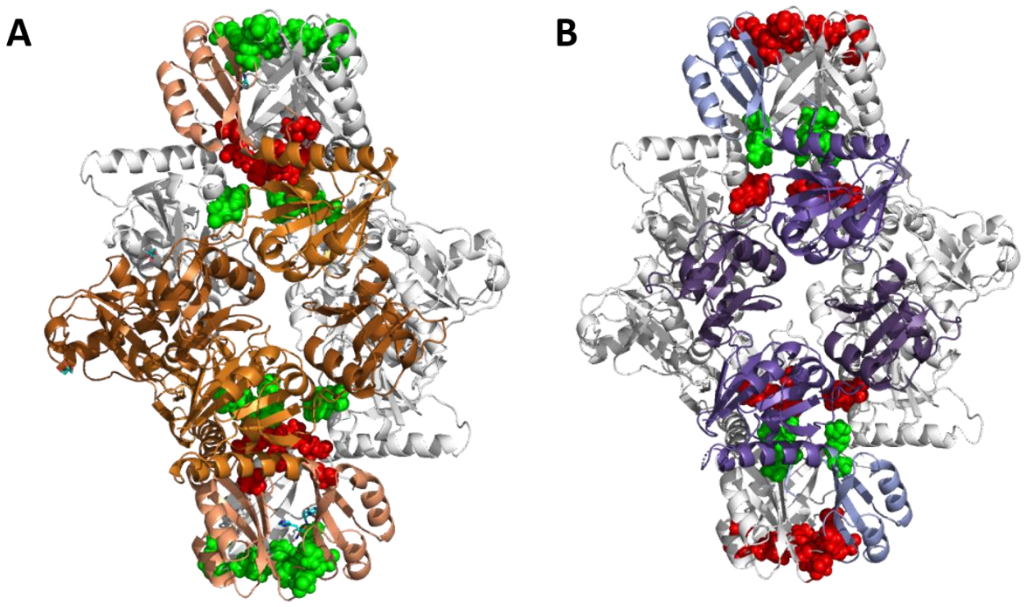

Figure 6.5. Illustrated H-Bonds formed (green spheres) and broken (red spheres) in *mtu*ATP-PRT$_{WT}$ (**A**) and *mtu*ATP-PRT$_{D258N}$ (**B**) between L-His bound states. Bonds formed or broken are defined as interactions that occur in greater or lesser propensity (10% frequency change compared to other allosteric states) as identified in MD simulations (**Table 6.4**).

Although crystallographic analysis and MD take two separate approaches to investigating changes to a protein; surprisingly, MD simulations identified H-bond changes in residues Arg27 and Glu18 that was also identified in crystal structure analysis (**Figure 5.19**). From MD simulations in *mtu*ATP-PRT$_{WT}$ (**Figure 6.6**) this bond is formed, as an inter-chain interaction. In Δ258N, upon binding of L-His, this H-bond is lost (**Figure 6.7**). Although this inter-chain interaction can be observed in **Figure 6.6** and **Figure 6.7**, in figures **Figure 6.3**, **Figure 6.4**, **Figure 6.5**, it appears as an intra-chain interaction, highlighting the additional information that can be obtained *via* MD simulations that is otherwise missed through examining crystallographic structures.

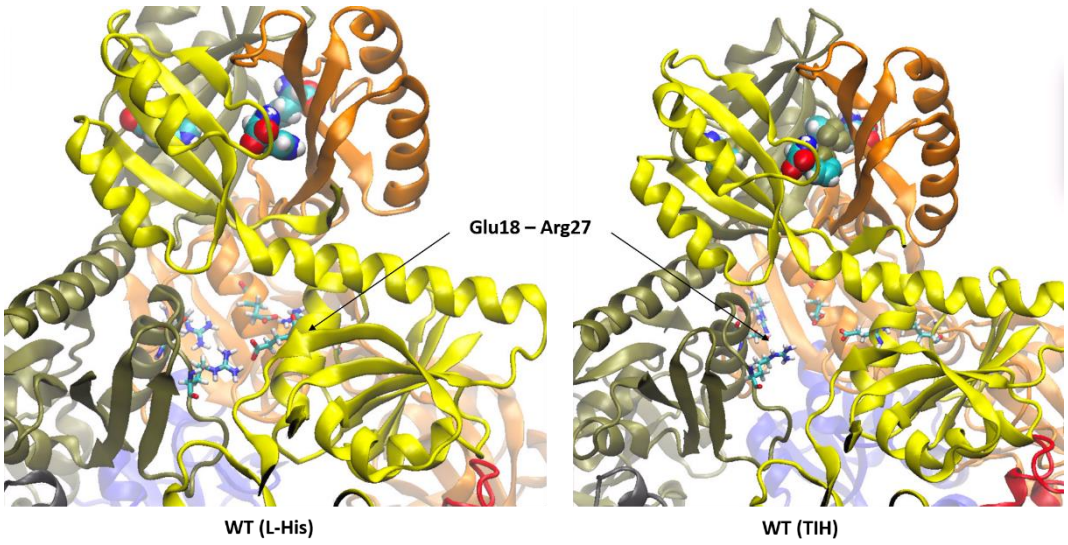

Figure 6.6. Glu18 – Arg 27 bond interaction captured from MD simulations in *mtu*ATP-PRT$_{WT}$. Glu18 and Arg27 (sticks) showing inter-chain interaction when L-His is bound that is lost with the binding of TIH in *mtu*ATP-PRT$_{WT}$.

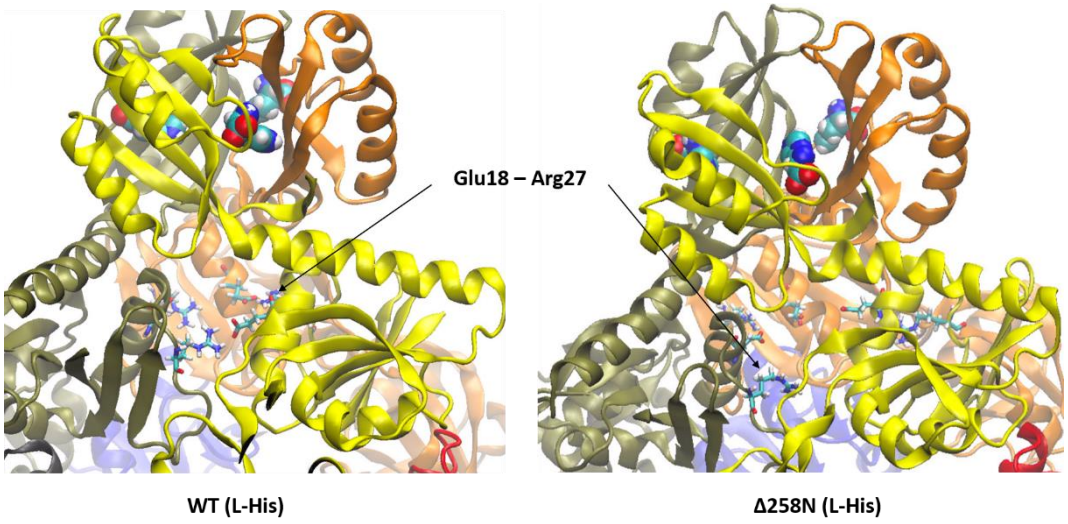

Figure 6.7. Glu18 – Arg 27 bond interaction captured from MD simulations in between *mtu*ATP-PRT varients. Glu18 and Arg27 (sticks) showing inter-chain interaction when L-His is bound in *mtu*ATP-PRT$_{WT}$. In Δ258N this interaction is lost with the binding of L-His.

6.3 Discussion

Although only preliminary, these MD results provide an additional suite of data that can be combined with kinetic and crystallographic data to investigate the mechanisms of allosteric regulation in *mtu*ATP-PRT. This section will discuss the

results of MD simulations and will be further expanded in Chapter 7 to bring together data from the multiple sources.

*mtu*ATP-PRT$_{D258N}$ with L-His bound had the longest equilibrated time compared to the other five simulations that ran during the course of the experiment (**Table 6.1**, **Figure 6.1**). This long equilibration period suggests the starting conformation of this structure may not be fully representative of the state of the enzyme during simulation. This is of course a function of the crystallographic structure representing the protein in a lattice, whereas the MD is modelling that of the solubilised protein in water. A longer equilibration period indicates a greater distance between the two states; the state of the crystal structure, and the state of the protein during simulation. It may also indicate a higher energy barrier between the two states that must be overcome. This structure was solved in the tensed state, like all other ligand-bound *mtu*ATP-PRT structures. However, due to the lengthy equilibration period there is the possibility that the crystallisation condition or refinement method have forced *mtu*ATP-PRT$_{D258N}$ with L-His in the tensed state, which not representative of the enzyme in solution. This disparity between crystallographic and simulation results highlights the constraints seen in crystallography, and enzyme structure may differ in aqueous states not captured by crystallography. Although *mtu*ATP-PRT$_{D258N}$ (L-His) had an increased equilibration time, data from all simulations were trimmed after the protein had equilibrated and a relaxed protein state was achieved, adding consistency to the analysis.

RMSD analysis of the backbones of each protein and allosteric state within the simulation are consistent with our understanding of flexibility changes. These, along with H-bond changes support the idea that variation in flexibility of an enzyme contributes to its catalysis under allosteric modulation. In *mtu*ATP-PRT$_{WT}$ we see moderate flexibility of the protein during simulation (**Figure 6.2**), that is lost during inhibition by L-His and regained upon allosteric activation by TIH. This supports the idea that tension is a function of reduced flexibility, and thus reduced flexibility of an enzyme system can manifest as reduced catalysis. Upon TIH binding, this overall flexibility was regained.

In the Δ258N data we see L-His binding having lesser variation to the flexibility of the protein backbone, indicating that although allosteric inhibitors are bound, the enzyme is still able to move and is not held in as rigid a state as in WT (L-His). The Δ258N variant does show reduced flexibility compared to the APO state, however the reduction in flexibility is not as severe as seen in WT (*i.e.* the reduction in the error bars is less pronounced). These RMSD data correspond to the kinetic data collected in (**Figure 4.12**). Kinetically, we see *mtu*ATP-PRT$_{D258N}$ shows continued activity whilst in the presence of L-His, indicating reduced sensitivity to the allosteric inhibitor. RMSD data compliment this, with reduced variation in RMSD between APO and L-His bound states observed in **Figure 6.2**. Furthermore, there is a localised region of bond breaking between the first and third domains (**Figure 6.5**, **C**), that is lesser than those bonds broken in L-His sensitive WT **(Figure 6.5**, **A**). This indicates more bonds are remaining in this L-His state, bringing it to a more similar state as the APO protein, perhaps contributing to the L-His insensitivity.

H-bond analysis from the simulations was able to identify a number of formed and broken bonds between each allosteric state (**Table 6.2**, **Table 6.3** and, **Table 6.4**). Interestingly, none of the significant H-bonds identified were involved with the mutated residue 258, which experimentally altered L-His response. Position 258 is located on the surface of Domain III, and thus highly solvent exposed. Many bond changes observed during allosteric ligand binding were around the C-terminus, surrounding this mutation site and involved in the overall domain movement that encompasses the tensed state. This aligns with literature and existing crystallographic data, along with crystallographic data collected in this research which shows the enzyme adopting a tensed state that features tightening of the C-terminus in favour over the trimeric interface and catalytic domains. Some of these C-terminus bonds were altered in Δ258N, and instead were formed closer to the loop region between domains I and III, near residue 258.

Although no H-bonds involved residue 258D or 258N, kinetic data clearly supports some involvement of residue 258D and 258N in allosteric inhibition response to L-His. The effect of mutation may lie in the backbone of the protein, or through interactions that are present at a low occupancy, and thus these subtle movements were cut off during this preliminary analysis. As this research encompasses a

preliminary view of MD simulations of *mtu*ATP-PRT, the current analysis is at a coarse level, however, does provide directions for further work. If the backbone drives the response to L-His then it would be pertinent to closely examine angles of the backbone around the allosteric site in great detail. Another suggestion is to adjust the cut off of occupancy and bond changes, particularly around critical sections of mutation and ligand binding in order to identify low occupancy bond changes or subtle interactions.

Although the response to L-His is different between WT and Δ258N variants, there was significant commonality seen in the bonds broken and formed between the two variants. Not only are the significant occupancy of bonds shared, but the distribution of bonds broken occur in a very restricted 3D location below the allosteric site (**Figure 6.4**, **A** and **C**). The broken bonds appear to reduce the coupling of the regulatory domain away from the main structure of the enzyme, which may be a source of reduced activity. **Figure 6.4 B** and **D** show bond formation between the two variants *mtu*ATP-PRT$_{WT}$ (**B**), and *mtu*ATP-PRT$_{D258N}$ (**D**). In L-His insensitive *mtu*ATP-PRT$_{D258N}$ there is a reduction in formed bonds around domains I and II which form the active site, compared to WT. There are also fewer bonds formed in domain III (allosteric domain). The source of histidine inhibition may be involved with the reduced bonding between the allosteric site and catalytic core. Effectively removing vibrational energy stored in the allosteric site from contributing to the catalysis.

Of particular interest was identification of an interaction between residues Glu18 and Arg27. These two residues were identified as key residues in (**Figure 5.19** and **Figure 5.20**) and this will be further discussed in Chapter 7 in context examining kinetic, crystallographic, and simulation data as a whole. MD simulations identified during L-His inhibition of WT protein Glu18 and Arg27 form an inter-chain interaction *via* a salt bridge. In WT APO and TIH structures this interaction is lost, opening up the structure and reducing the interaction between these chains. In Δ258N with L-His bound this interaction is also lost, accompanying lost inhibition to L-His. Thus, this interaction is observed in the MD only in instances where we know inhibition to be occurring from our kinetics data. With this loss of interaction, we suggest Glu18 and Arg27 form a significant interaction in allosteric inhibition that may increase the stability of the hexamer, and thus reduce the flexibility of the

enzyme, and the catalytic capacity that is dependent on this flexibility. The notion of bond formation in L-His bound states implies decreased flexibility, and this builds on other MD simulation data suggesting reduced flexibility in L-His bound enzymes.

In *mtu*ATP-PRT$_{WT}$ (**Figure 6.6**) Glu18 and Arg27 bond is formed, as an inter-chain interaction. In Δ258N upon binding of L-His, this H-bond is lost (**Figure 6.7**). This was also observed in crystal data (**Figure 5.19** and **Figure 5.20**), however, the interaction was observed as an intra-chain interaction tightening domain I near the active site (**Figure 5.19**). Whilst it is exciting to see a cross over from crystallographic and MD simulation data, this interaction is isolated and how this links to the overall allosteric communication is unknown. However, identifying the same isolated residue movement in independent static crystallography analysis and aqueous simulation suggests this interaction to be of importance in explaining enzyme changes in allosterically active and inhibited states.

Unfortunately, although changes between *mtu*ATP-PRT variants and allosteric states were observed, preliminary MD simulations did not provide conclusive evidence for how the allosteric ligands elicit overall structural changes, nor how TIH elicits an opposing catalytic effect to L-His. However, it has highlighted regions for further investigation. There was an intriguing lack of difference around the regions of the protein actually altered (residue 258, nor around the ligand binding site). Further work increasing replicate numbers, and looking closer at mutation sites, lowering thresholds for H-bond occupancy, and observing the backbone angles near active and allosteric sites is necessary. It would be of interest to see what H-bonds are altered in Δ258A as these may assist in narrowing down key residues driven by each mutation and ligand.

6.4 Conclusion

Through preliminary MD simulations using two *mtu*ATP-PRT variants in three allosteric states we have identified one key residue interaction (Glu18 - Arg27) that exists in L-His inhibited WT states, stabilising the enzyme in a tensed, less active state. This residue interaction was not seen in the Δ258N (L-His) mutant, which is kinetically insensitive to L-His, suggesting this could be a residue switch that drives

the tensed inactive form of *mtu*ATP-PRT. RMSD analysis of the protein backbones show in simulated L-His inhibited structures have reduced RMSD and thus are more rigid compared to structures with activating ligands or no allosteric regulation. Future work of collecting simulation data for Δ258A in different allosteric states would help identify key residues in allosteric modulation in *mtu*ATP-PRT and assist in unravelling the pathways this enzyme is using for allosteric signal transduction.

Chapter 7
Discussion

Enzymes are the driving catalyst of countless reactions within cells, drastically increasing chemical reaction rates (Fersht, 1999; Wolfenden & Snider, 2001; Daniel & Danson, 2010). These catalytic effects impact a diverse number of reactions, with intricate systems of regulation at many levels acting to control enzyme rates and fine-tune metabolism. Allostery is a mechanism for feedback and enzyme regulation which describes modulation of enzymes where the modulating ligand (generally bearing no similarity to substrate) binds to a distinct site, distal to the active site of the enzyme and exerts an alteration to enzyme activity (Motlagh *et al.*, 2014; Liu & Nussinov, 2016). Various conformational and dynamic models have been developed to explain the mechanism of allosteric regulation in enzymes (Perutz *et al.*, 1960; Monod *et al.*, 1965; Koshland *et al.*, 1966; Weber, 1972; Cooper & Dryden, 1984; Jaffe, 2005; Hilser *et al.*, 2012). ATP-PRT is an interesting case, being an enzyme existing in two unique isoforms in different organisms with both being allosterically regulated by L-His. Some ATP-PRT are additionally allosterically activated, thus ATP-PRT is an ideal model for investigation into allosteric mechanism *via* kinetic and structural analysis, and analysis using established thermodynamic principles.

Investigation of the driving mechanism of allostery in ATP-PRT requires determination of how allosteric mechanisms alter the structure of ATP-PRT, how catalysis alters in response to allosteric ligands, how changes to catalysis alter the energy landscape of the reaction, and what communication pathways or key residues are involved with allosteric communication.

Experimentally, *mtu*ATP-PRT$_{WT}$ showed increased catalysis in the presence of the allosteric activator TIH, and decreased catalysis in the presence of the allosteric inhibitor L-His (**Figure 4.12 A**). The increased catalysis by TIH was accompanied by a near similar T_{opt} to APO, and a slightly reduced $\Delta C_p^{\ddagger}{}_{T0}$, $\Delta H^{\ddagger}{}_{T0}$, and a more negative $\Delta S^{\ddagger}{}_{T0}$. The temperature rate relationship between APO and TIH activated *mtu*ATP-PRT$_{WT}$ was very similar across the collected temperature range. In

contrast, L-His elicited reduced activity to *mtu*ATP-PRT$_{WT}$ in physiologically relevant levels, with the inhibitory effect lost at lower and upper temperature extremes.

Michaelis-Menten plots with L-His bound showed an increased K_M for substrate PRPP in response to histidine at 298 and 313 K. An increased K_M can be anticipated in allosteric inhibition, as an indicator of lower affinity for substrate. At physiological temperature (313 K), the K_M for PRPP was increased in the presence of inhibitor L-His. This decreased affinity for substrate (in the presence of L-His) supports the work of Jiao *et al.* (2019) who observed tightening in the active site of *cje*ATP-PRT associated with L-His binding causing a steric clash of substrates, thus altering substrate binding and reducing catalysis. However, allosteric activation by TIH elicits a near identical protein structure which also includes the steric clash (Pisco *et al.*, 2017), which then appears to contradict the proposed mechanism of steric clashes driving allosteric inhibition.

In addition to assessing the thermodynamic profile of *mtu*ATP-PRT$_{WT}$ in response to allosteric modulation, residue 258 was targeted to attempt to alter sensitivity to L-His in *mtu*ATP-PRT. Two alternate residues were substituted at this position; *mtu*ATP-PRT$_{D258A}$ and *mtu*ATP-PRT$_{D258N.}$ This site directed mutation altered the response to L-His, with *mtu*ATP-PRT$_{D258A}$ showing increased sensitivity, and *mtu*ATP-PRT$_{D258N}$ showing reduced sensitivity to L-His. Interestingly, both mutations retained sensitivity to TIH, with *mtu*ATP-PRT$_{D258A}$ showing exaggerated sensitivity at some temperatures. Whilst residue 258 is not directly involved in binding to the allosteric ligands L-His or TIH, the change in sensitivity to the allosteric inhibitor ligand certainly implicates the role of residue 258 in allosteric inhibition. With residue 258 not resulting in alterations to allosteric activation by TIH, this opens the idea that L-His and TIH elicit allosteric effect *via* different allosteric communication pathways across the 40 Å distance from the regulatory domain to the active site.

The MMRT fits from T_{opt} assays (**Figure 4.12, Table 4.1,** and **Table 4.2**) show a clear enthalpy-entropy trade-off that exists between allosteric ligand modulated enzyme rates (**Figure 4.13**). *mtu*ATP-PRT$_{WT}$ serves as an archetype of this relationship, with activated, APO, and inhibited *mtu*ATP-PRT$_{WT}$ behaving akin to

psychrophilic, mesophilic, and thermophilic enzymes in the observed changes to the energy barrier to the reaction (Arcus & Mulholland, 2020).

In the case of L-His inhibited (the analogous thermophilic enzyme), the energy barrier of the reaction is increased *via* increased $\Delta H^{\ddagger}$. In response to the increased $\Delta H^{\ddagger}$ barrier, the $\Delta S^{\ddagger}$ barrier of the system is reduced, along with a reduction to $\Delta C_p^{\ddagger}$ signifying reduced fluctuations of the enzyme in the enzyme-substrate complex. Comparatively, activation by TIH shows a decreased $\Delta H^{\ddagger}$ barrier thus a reduction to the energy barrier of the reaction. Like in inhibition by L-His, the alteration to $\Delta H^{\ddagger}$ must be traded for an increase to the $\Delta S^{\ddagger}$ barrier which results in altered $\Delta C_p^{\ddagger}$. Examining relaxed versus tensed states of *mtu*ATP-PRT$_{WT}$ and the number of hydrogen bonds associated with achieving the tensed enzyme state associated with allosteric inhibition we can see how the enthalpy of the system and respective flexibility of the enzyme changes in response to L-His. There is evidence of additional bonds formed in L-His bound states compared to APO or TIH bound states. These additional bonds may function in the tensing of the enzyme, thus reducing the flexibility of the overall structure and increasing the enthalpic barrier.

When mutations were made to residue 258 of *mtu*ATP-PRT these resulted in alteration of kinetic and allosteric behaviour. In *mtu*ATP-PRT$_{D258N}$, inhibition by L-His was lessened, and the associated $\Delta H^{\ddagger}_{T0}$ in the presence of L-His was decreased (compared to both *mtu*ATP-PRT$_{D258N}$ APO, and *mtu*ATP-PRT$_{WT}$ L-His inhibited), thus reducing the energy barrier of the reaction. In contrast *mtu*ATP-PRT$_{D258A}$ showed significant inhibition by L-His, this inhibition was reflected in the significantly larger $\Delta H^{\ddagger}_{T0,}$ values again, indicating an increased energy barrier.

To investigate structural changes to the enzyme that may alter bond formation and thus changes to the enthalpy barrier, crystal structures were collected in APO, and ligand bound states for *mtu*ATP-PRT$_{D258A}$ and *mtu*ATP-PRT$_{D258N}$ (Chapter 5). The APO structures are thought to represent the relaxed form and align well with *mtu*ATP-PRT$_{WT}$ structure previously reported by Cho *et al.* (2003). The ligand bound structures also aligned with previously published *mtu*ATP-PRT$_{WT}$ structures by Pisco *et al.* (2017), indicating both L-His and TIH elicit the tensed conformation, even in enzymes with reduced sensitivity to the ligand.

These complete structures were used in H-bond analysis, which extracted each unique bond between APO and ligand bound states for each enzyme (Section 5.2.9). Both allosteric modulators for all variants propagated increased H-bond interactions in the allosteric domain, which relate to the movement into the tensed state. As the H-bond analysis is derived from the solved crystal structures, it only serves to represent the difference between the static crystal structures collected, and thus may provide a generalised view of allosteric change. We acknowledge that different protein purification and crystallographic conditions could contribute to differences in bond interactions that are not relevant to the interactions driving the allosteric effect. In addition, MD simulations showed increased equilibration times for ligand bound structures (particularly the two TIH bound structures), indicating at least one of these structures was solved in a state that may not be representative of the enzyme in solution. As a consequence, H-bond results identified from crystal structures should be used as a method to detect areas of change between protein structures, but may not be suitable in identifying specific changes that occur during the enzymatic reaction.

Despite the potential caveats for our H-bond analysis, the bonding in structural data shows a high linkage between the chains of the oligomer. The increasing number of inter-chain bonds seen in the ligand bound states illustrate the degree of cooperativity involved in allosteric response. Cooperative effect by L-His has been suggested by Mittelstädt (2015) in *cje*ATP-PRT. Michaelis-Menten plots (**Section 4.2.4**) did not indicate this, but primary isothermal titration calorimetry (**Table 8.1** and **Figure A.6.1**) indicate one L-His binding per hexamer. Suggesting the binding of an allosteric molecule to a single allosteric site stimulates a global change to the enzyme structure.

These same solved structures were used in MD simulations (Chapter 6). In the case of *mtu*ATP-PRT$_{D258N,}$ the variant with reduced L-His sensitivity, the crystal structure was solved in a tensed state, like *mtu*ATP-PRT$_{WT}$. Upon simulation, there was a significant increase in equilibration time. The increased equilibration time suggests the crystal structure may not be representative of the structure in solution (Walton & VanVliet, 2006), and perhaps this residue alteration has altered the structural state or the frequency in which it engages in the tensed phase. If *mtu*ATP-PRT$_{D258N}$ is in fact not in the tensed state (as observed by crystallography), but

instead is in a relaxed state containing less constraining bonds, this may explain the reduced sensitivity to L-His for this variant.

MD analysis identified reduced bond formation in the C-terminus of *mtu*ATP-PRT_{D258N} in the presence of L-His compared to *mtu*ATP-PRT_{WT} under the same conditions. Instead bond formation was focused in the hinge region between Domain I and III (**Figure 6.5**). This indicates bond formation around the regulatory domain (which is characteristic of the tight hexamer orientation of the T-state ATP-PRT) is reduced in the L-His insensitive Δ258N variant. This indicates fewer bonds are forming around the regulatory domain, thus the enzyme may not be entering the T-state as we know it, consistent with the altered L-His response in this variant.

Both MD simulations and comparisons of crystal structures identified the movement of residues Glu18 and Arg27 of domain I between APO and TIH states. This interaction occurs as an intra-chain interaction in crystal structures, and inter-chain interaction in MD simulations. This is the first observation of altered residues near the active site in response to allosteric activation. Through crystallographic analysis it appears Glu18 and Arg27 interact under allosteric activation by TIH. This interaction is intra-chain in static crystal structures, and bridged across a section of the active site in domain I. Initially, this was thought to add stability to the active site and increase the catalytic effect of the mutant. However, MD simulations observed this bond was formed in an inter-chain fashion under allosteric inhibition by L-His, and this interaction was relieved in the APO and TIH bound states. An inter-chain interaction between these residues present would certainly tighten the structure of the enzyme and reduce the flexibility, thus we might expect an increase in the enthalpy barrier in the presence of L-His.

Following the identification of Glu18 and Arg27 interactions in L-His bound structures, it would be pertinent to inspect whether this interaction is an artefact of protein crystal structure or a fundamental interaction in allosteric activation. Firstly mutating each residue to Alanine to see if this elicits a reduced effect of allosteric inhibition. Mutation of each residue Glu18 and Arg27 to cysteine may promote cysteine bridge formation. Keeping in mind the distance between the inter-chain interaction at 9.8 Å may be too great of a distance to cover, and this could instead facilitate the formation of an intra-chain disulphide interaction at 5.4 Å distance.

However, if allosteric inhibition by L-His alters the position of residue 18 and 27, and if this disulphide bond remained present, we could expect inhibition to the enzyme without the presence of the inhibitor, therefore proving the interaction identified in MD simulations. If this E18C R27C mutant showed increased catalysis in the APO form it may suggest these residues are instead involved in allosteric activation by TIH, as observed in crystallographic analysis.

Through the information obtained in T_{opt} and MMRT analysis, crystallography, and MD simulations we propose an enthalpy-entropy trade-off between allosterically modulated states that controls the temperature rate relationship. For allosteric inhibition by L-His, the formation of an inter-chain Glu18 Arg27 interaction creates additional bonds in *mtu*ATP-PRT that reduce the flexibility of the enzyme and thus alter the energy barrier of the reaction.

Chapter 8

Future recommendations and experiments

Isothermal titration calorimetry (ITC) across various temperatures would allow for the determination of binding constants across temperature and ATP-PRT variant and also allows the calculation of ΔC_p for binding from the temperature-dependence of ΔH. In preliminary ITC data using *mtu*ATP-PRT_{WT} with the addition of L-His, we observed a shift from endothermic to exothermic binding between 20 and 37 °C (**Table 8.1**; full graph in Appendix 6). Both 20 and 37 °C showed similar K_D values of 11.3 and 11.6 µM for L-His binding.

Table 8.1. Isothermal titration calorimetry values for binding of L-His to *mtu*ATP-PRT_{WT}.

Temperature	K_D (µM)	N	ΔH (kJ/mol)	ΔS (J/mol. K)
20 °C	11.3	0.16	16.9	149.9
27 °C	250.0	0.18	100.0	402.0
37 °C	11.6	0.39	4.27	105.2

There is an endothermic to exothermic binding shift between 20 and 37 °C, and we suggest 27 °C is possibly in the middle of this shift, resulting in the unusual data collected. It would certainly be beneficial to collect additional binding data on both L-His and TIH to each of the ATP-PRT variants, across a range of temperatures. This data would then inform how allosteric ligand binding sensitivity changed across the temperature profile, and how this correlated with kinetic data already collected.

MD simulations showed increased equilibration on *mtu*ATP-PRT_{D258N} in the presence of L-His. Additional simulations using the collection of *mtu*ATP-PRT_{D258A} would be valuable in identifying whether this equilibration delay is a function of crystallography or whether the delay was indicative of the tensed state not representing the true structure of L-His inhibited *mtu*ATP-$PRT_{D258N.}$ Furthermore, with *mtu*ATP-PRT_{D258A} remaining sensitive to both L-His and TIH

additional MD simulations, and H bond analysis may be useful in identifying key residues involved in inhibition.

These enzymes and their allosteric response clearly varies across a temperature range. Running MD simulations across various temperature ranges may allow us to see if temperature plays a role in the enzyme conformation, and what residues or bonds change in different temperature states. It would be particularly interesting to collect simulation data at temperatures that encompassed the crossover of relative rates observed in the T_{opt} data (**Figure 4.16**).

Both MD and structural analysis identified an alteration to residues Glu18 and Arg27 in the presence of TIH. Identifying the same residues *via* two independent methods holds great promise these may be functioning alterations involved in allosteric activation. If each were mutated to cysteine or truncated, we may see altered catalysis in *mtu*ATP-PRT without the addition of allosteric ligand. For disulphide bond formation, how would the addition of reducing agents to interrupt this interaction to investigate if this then reduces rates. In the instance of truncated mutants how would this alter the enzymes catalytic rate? Would this this alter the enzymes ability to enter the T-state, and would this be reflected in the equilibration times during subsequent MD simulations?

Structural analysis performed during this thesis did not identify dramatic changes between ligand bound structures or ATP-PRT variants that could be expected with the large shift to catalytic effect observed in kinetic data. All crystal structures collected at cryogenic conditions not representative of protein conformations at room or physiological temperatures, and the collection of data at cryogenic temperatures is thought to alter the protein backbone (Fraser *et al.*, 2011). Thus the move to room-temperature crystallography may provide additional information on the protein backbone movement, and elucidate the protein dynamics driving allosteric regulation. Particularly interesting in the case of L-His and TIH bound structures which solve as near indistinguishable structures despite the dramatic effect the ligand has on catalysis.

There has been work done by Subramanian *et al.* (2021) to apply MMRT to a single molecule level to deeply investigate conformational change and turnover of an

enzyme. After immobilising the enzyme, and rinsing away remaining enzymes, data is able to be collected and the sensitivity of data collection is able to identify conformational fluctuation of enzyme dynamics. It is possible this method is suitable to measuring the conformational changes associated with allosteric modulation, and the high sensitivity of the method may allow us to discern any changes between L-His and TIH bound systems that were suggested by MD simulations.

References

Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H., & Adams, P. D. (2012). Towards automated crystallographic structure refinement with phenix. refine. *Acta Crystallographica Section D: Biological Crystallography, 68*(4), 352-367.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control, 19*(6), 716-723.

Aklujkar, M. (2011). Two ATP phosphoribosyltransferase isozymes of Geobacter sulfurreducens contribute to growth in the presence or absence of histidine and under nitrogen fixation conditions. *Canadian journal of microbiology, 57*(7), 547-558.

Alm, E. J., Huang, K. H., Price, M. N., Koche, R. P., Keller, K., Dubchak, I. L., & Arkin, A. P. (2005). The MicrobesOnline Web site for comparative genomics. *Genome research, 15*(7), 1015-1022.

Ames, B. N., Martin, R. G., & Garry, B. J. (1961). The First Step of Histidine Biosynthesis. *Journal of Biological Chemistry, 236*(7), 2019-2026.

Arcus, V. L., & Mulholland, A. J. (2020). Temperature, dynamics, and enzyme-catalyzed reaction rates. *Annual review of biophysics, 49*, 163-180.

Arcus, V. L., Prentice, E. J., Hobbs, J. K., Mulholland, A. J., Van der Kamp, M. W., Pudney, C. R., Parker, E. J., & Schipper, L. A. (2016). On the temperature dependence of enzyme-catalyzed rates. *Biochemistry, 55*(12), 1681-1688.

Arcus, V. L., & Pudney, C. R. (2015). Change in heat capacity accurately predicts vibrational coupling in enzyme catalyzed reactions. *FEBS letters, 589*(17), 2200-2206.

Armougom, F., Moretti, S., Poirot, O., Audic, S., Dumas, P., Schaeli, B., Keduas, V., & Notredame, C. (2006). Expresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee. *Nucleic Acids Research, 34*(suppl_2), W604-W608.

Battistuzzi, F. U., Feijao, A., & Hedges, S. B. (2004). A genomic timescale of prokaryote evolution: insights into the origin of methanogenesis, phototrophy, and the colonization of land. *BMC Evolutionary Biology, 4*(1), 44.

Blum, M., Chang, H.-Y., Chuguransky, S., Grego, T., Kandasaamy, S., Mitchell, A., Nuka, G., Paysan-Lafosse, T., Qureshi, M., Raj, S., Richardson, L., Salazar, G. A., Williams, L., Bork, P., Bridge, A., Gough, J., Haft, D. H., Letunic, I., Marchler-Bauer, A., Mi, H., Natale, D. A., Necci, M., Orengo, C. A., Pandurangan, A. P., Rivoire, C., Sigrist, C. J. A., Sillitoe, I., Thanki, N., Thomas, P. D., Tosatto, S. C. E., Wu, C. H., Bateman, A., & Finn, R.

D. (2020). The InterPro protein families and domains database: 20 years on. *Nucleic Acids Research, 49*(D1), D344-D354.

Bond, J. P., & Francklyn, C. (2000). Proteobacterial histidine-biosynthetic pathways are paraphyletic. *Journal of molecular evolution, 50*(4), 339-347.

Bowers, K. J., Chow, D. E., Xu, H., Dror, R. O., Eastwood, M. P., Gregersen, B. A., Klepeis, J. L., Kolossvary, I., Moraes, M. A., & Sacerdoti, F. D. (2006). Scalable algorithms for molecular dynamics simulations on commodity clusters. In *SC'06: Proceedings of the 2006 ACM/IEEE Conference on Supercomputing* (pp. 43-43): IEEE.

Brenner, M., & Ames, B. N. (1971). CHAPTER 11 - The Histidine Operon and Its Regulation. In H. J. Vogel (Ed.), *Metabolic Regulation (Third Edition)* (https://doi.org/10.1016/B978-0-12-299255-1.50018-3pp. 349-387). Academic Press. https://doi.org/10.1016/B978-0-12-299255-1.50018-3.

Bubić, A., Mrnjavac, N., Stuparević, I., Łyczek, M., Wielgus-Kutrowska, B., Bzowska, A., Luić, M., & Leščić Ašler, I. (2018). In the quest for new targets for pathogen eradication: the adenylosuccinate synthetase from the bacterium Helicobacter pylori. *Journal of enzyme inhibition and medicinal chemistry, 33*(1), 1405-1414.

Cambau, E., & Drancourt, M. (2014). Steps towards the discovery of Mycobacterium tuberculosis by Robert Koch, 1882. *Clinical Microbiology and Infection, 20*(3), 196-201.

Carmel-Harel, O., & Storz, G. (2000). Roles of the glutathione-and thioredoxin-dependent reduction systems in the Escherichia coli and Saccharomyces cerevisiae responses to oxidative stress. *Annual Reviews in Microbiology, 54*(1), 439-461.

Catlin, B. W. (1973). Nutritional Profiles of Neisseria gonorrhoeae, Neisseria meningitidis, and Neisseria lactamica in Chemically Defined Media and the Use of Growth Requirements for Gonococcal Typing. *The Journal of Infectious Diseases, 128*(2), 178-194.

Celej, M. S., Montich, G. G., & Fidelio, G. D. (2003). Protein stability induced by ligand binding correlates with changes in protein flexibility. *Protein science : a publication of the Protein Society, 12*(7), 1496-1506.

Champagne, K. S., Piscitelli, E., & Francklyn, C. S. (2006). Substrate recognition by the hetero-octameric ATP phosphoribosyltransferase from Lactococcus lactis. *Biochemistry, 45*(50), 14933-14943.

Champagne, K. S., Sissler, M., Larrabee, Y., Doublié, S., & Francklyn, C. S. (2005). Activation of the Hetero-octameric ATP Phosphoribosyl Transferase through Subunit Interface Rearrangement by a tRNA Synthetase Paralog. *Journal of Biological Chemistry, 280*(40), 34096-34104.

Changeux, J.-P. (2012). Allostery and the Monod-Wyman-Changeux model after 50 years. *Annual review of biophysics, 41*, 103-133.

Chipman, D. M., & Shaanan, B. (2001). The ACT domain family. *Current opinion in structural biology, 11*(6), 694-700.

Cho, Y., Sharma, V., & Sacchettini, J. C. (2003). Crystal Structure of ATP Phosphoribosyltransferase from Mycobacterium tuberculosis. *Journal of Biological Chemistry, 278*(10), 8333-8339.

Clavel, T., Lepage, P., & Charrier, C. (2014). The Family Coriobacteriaceae. In E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt & F. Thompson (Eds.), *The Prokaryotes: Actinobacteria* (10.1007/978-3-642-30138-4_343pp. 201-238). Berlin, Heidelberg: Springer Berlin Heidelberg. 10.1007/978-3-642-30138-4_343.

Colombo, M. F., Rau, D. C., & Parsegian, V. A. (1992). Protein solvation in allosteric regulation: a water effect on hemoglobin. *Science, 256*(5057), 655-9.

Cooper, A., & Dryden, D. T. (1984). Allostery without conformational change. A plausible model. *Eur Biophys J, 11*(2), 103-9.

Cross, P. J., Allison, T. M., Dobson, R. C. J., Jameson, G. B., & Parker, E. J. (2013). Engineering allosteric control to an unregulated enzyme by transfer of a regulatory domain. *Proceedings of the National Academy of Sciences, 110*(6), 2111-2116.

Cui, Q., & Karplus, M. (2008). Allostery and cooperativity revisited. *Protein science, 17*(8), 1295-1307.

D'Amico, S., Sohier, J., & Feller, G. (2006). Kinetics and energetics of ligand binding determined by microcalorimetry: insights into active site mobility in a psychrophilic α-amylase. *Journal of molecular biology, 358*(5), 1296-1304.

Daniel, R. M., & Danson, M. J. (2010). A new understanding of how temperature affects the catalytic activity of enzymes. *Trends in Biochemical Sciences, 35*(10), 584-591.

Daniel, R. M., Danson, M. J., & Eisenthal, R. (2001). The temperature optima of enzymes: a new perspective on an old phenomenon. *Trends in biochemical sciences, 26*(4), 223-225.

del Sol, A., Tsai, C.-J., Ma, B., & Nussinov, R. (2009). The origin of allosteric functional modulation: multiple pre-existing pathways. *Structure (London, England : 1993), 17*(8), 1042-1050.

Delorme, C., Ehrlich, S. D., & Renault, P. (1999). Regulation of Expression of the *Lactococcus lactis* Histidine Operon. *Journal of Bacteriology, 181*(7), 2026-2037.

Dodson, E. J. (1991). Baverage.

Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. (2010). Features and development of Coot. *Acta Crystallographica Section D: Biological Crystallography, 66*(4), 486-501.

Evans, P. R. (2011). An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallographica Section D: Biological Crystallography, 67*(4), 282-292.

Evans, P. R., & Murshudov, G. N. (2013). How good are my data and what is the resolution? *Acta Crystallographica Section D: Biological Crystallography, 69*(7), 1204-1214.

Eyring, H. (1931). The energy of activation for bimolecular reactions involving hydrogen and the halogens, according to the quantum mechanics. *Journal of the American Chemical society, 53*(7), 2537-2549.

Fani, R., Brilli, M., & Lio, P. (2005). The origin and evolution of operons: the piecewise building of the proteobacterial histidine operon. *Journal of molecular evolution, 60*(3), 378-390.

Feller, G., & Gerday, C. (2003). Psychrophilic enzymes: hot topics in cold adaptation. *Nature reviews microbiology, 1*(3), 200-208.

Ferreira, C. M., Pinto, I. S., Soares, E. V., & Soares, H. M. (2015). (Un) suitability of the use of pH buffers in biological, biochemical and environmental studies and their interaction with metal ions–a review. *Rsc Advances, 5*(39), 30989-31003.

Fersht, A. (1999). *Structure and mechanism in protein science: a guide to enzyme catalysis and protein folding*. Macmillan.

Fisher, G., Thomson, C. M., Stroek, R., Czekster, C. M., Hirschi, J. S., & da Silva, R. G. (2018). Allosteric activation shifts the rate-limiting step in a short-form ATP phosphoribosyltransferase. *Biochemistry, 57*(29), 4357-4367.

Fondi, M., Emiliani, G., Liò, P., Gribaldo, S., & Fani, R. (2009). The Evolution of Histidine Biosynthesis in Archaea: Insights into the his Genes Structure and Organization in LUCA. *Journal of Molecular Evolution, 69*(5), 512.

Fraser, J. S., van den Bedem, H., Samelson, A. J., Lang, P. T., Holton, J. M., Echols, N., & Alber, T. (2011). Accessing protein conformational ensembles using room-temperature X-ray crystallography. *Proceedings of the National Academy of Sciences, 108*(39), 16247.

Gallagher, L. A., Bailey, J., & Manoil, C. (2020). Ranking essential bacterial processes by speed of mutant death. *Proceedings of the National Academy of Sciences*, 10.1073/pnas.2001507117, 202001507.

Gandhi, N. R., Nunn, P., Dheda, K., Schaaf, H. S., Zignol, M., Van Soolingen, D., Jensen, P., & Bayona, J. (2010). Multidrug-resistant and extensively drug-resistant tuberculosis: a threat to global control of tuberculosis. *The Lancet, 375*(9728), 1830-1843.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic biology, 59*(3), 307-321.

Hamill, C. J. (2020). *Understanding allosteric enzyme regulation using macromolecular rate theory.* The University of Waikato.

Hilser, V. J., Wrabl, J. O., & Motlagh, H. N. (2012). Structural and energetic basis of allostery. *Annu Rev Biophys, 41*, 585-609.

Hobbs, J. K., Jiao, W., Easter, A. D., Parker, E. J., Schipper, L. A., & Arcus, V. L. (2013). Change in Heat Capacity for Enzyme Catalysis Determines Temperature Dependence of Enzyme Catalyzed Rates. *ACS Chemical Biology, 8*(11), 2388-2393.

Jaffe, E. K. (2005). Morpheeins--a new structural paradigm for allosteric regulation. *Trends Biochem Sci, 30*(9), 490-7.

Jaffe, E. K. (2016). The Remarkable Character of Porphobilinogen Synthase. *Accounts of chemical research, 49*(11), 2509-2517.

Jiao, W., Mittelstädt, G., Moggré, G.-J., & Parker, E. J. (2019). Hinge Twists and Population Shifts Deliver Regulated Catalysis for ATP-PRT in Histidine Biosynthesis. *Biophysical Journal, 116*(10), 1887-1897.

Jones, K. M., Lloret, J., Daniele, J. R., & Walker, G. C. (2007). The type IV secretion system of Sinorhizobium meliloti strain 1021 is required for conjugation but not for intracellular symbiosis. *Journal of bacteriology, 189*(5), 2133-2138.

Justi, R., & Gilbert, J. K. (1999). History and Philosophy of Science through Models: The Case of Chemical Kinetics. *Science & Education, 8*(3), 287-307.

Kabsch, W. (2010). Integration, scaling, space-group assignment and post-refinement. *Acta Crystallographica Section D: Biological Crystallography, 66*(2), 133-144.

Kamerlin, S. C. L., & Warshel, A. (2010). At the dawn of the 21st century: Is dynamics the missing link for understanding enzyme catalysis? *Proteins, 78*(6), 1339-1375.

Kantardjieff, K. A., & Rupp, B. (2003). Matthews coefficient probabilities: improved estimates for unit cell contents of proteins, DNA, and protein–nucleic acid complex crystals. *Protein Science, 12*(9), 1865-1871.

Kleeman, J. E., & Parsons, S. M. (1976). Reverse direction substrate kinetics and inhibition studies on the first enzyme of histidine biosynthesis, adenosine triphosphate phosphoribosyltransferase. *Arch Biochem Biophys, 175*(2), 687-93.

Koshland, D. E., Jr., Némethy, G., & Filmer, D. (1966). Comparison of experimental binding data and theoretical models in proteins containing subunits. *Biochemistry, 5*(1), 365-85.

Koul, A., Arnoult, E., Lounis, N., Guillemont, J., & Andries, K. (2011). The challenge of new drug discovery for tuberculosis. *Nature, 469*(7331), 483-490.

Kulis-Horn, R. K., Persicke, M., & Kalinowski, J. (2014). Histidine biosynthesis, its regulation and biotechnological application in Corynebacterium glutamicum. *Microbial biotechnology, 7*(1), 5-25.

Kulis-Horn, R. K., Persicke, M., & Kalinowski, J. (2015). Corynebacterium glutamicum ATP-phosphoribosyl transferases suitable for l-histidine production – Strategies for the elimination of feedback inhibition. *Journal of Biotechnology, 206*, 26-37.

Lang, E. J., Cross, P. J., Mittelstädt, G., Jameson, G. B., & Parker, E. J. (2014). Allosteric ACTion: the varied ACT domains regulating enzymes of amino-acid metabolism. *Current opinion in structural biology, 29*, 102-111.

Laskowski, R. A., Hutchinson, E. G., Michie, A. D., Wallace, A. C., Jones, M. L., & Thornton, J. M. (1997). PDBsum: a Web-based database of summaries and analyses of all PDB structures. *Trends in biochemical sciences, 22*(12), 488-490.

Li, X., Chen, Y., Lu, S., Huang, Z., Liu, X., Wang, Q., Shi, T., & Zhang, J. (2013). Toward an understanding of the sequence and structural basis of allosteric proteins. *J Mol Graph Model, 40*, 30-9.

Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L.-W., Jain, S., & McCoy, A. J. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallographica Section D: Structural Biology, 75*(10), 861-877.

Liebschner, D., Afonine, P. V., Moriarty, N. W., Poon, B. K., Sobolev, O. V., Terwilliger, T. C., & Adams, P. D. (2017). Polder maps: improving OMIT maps by excluding bulk solvent. *Acta Crystallographica Section D: Structural Biology, 73*(2), 148-157.

Liu, J., & Nussinov, R. (2016). Allostery: an overview of its history, concepts, methods, and applications. *PLoS computational biology, 12*(6), e1004966.

Livingstone, E. K., Mittelstädt, G., Given, F. M., & Parker, E. J. (2016). Independent catalysis of the short form HisG from Lactococcus lactis. *FEBS letters, 590*(16), 2603-2610.

Lo, M.-C., Aulabaugh, A., Jin, G., Cowling, R., Bard, J., Malamas, M., & Ellestad, G. (2004). Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. *Analytical biochemistry, 332*(1), 153-159.

Lu, S., Li, S., & Zhang, J. (2014). Harnessing allostery: a novel approach to drug discovery. *Medicinal research reviews, 34*(6), 1242-1285.

Malykh, E. A., Butov, I. A., Ravcheeva, A. B., Krylov, A. A., Mashko, S. V., & Stoynova, N. V. (2018). Specific features of l-histidine production by Escherichia coli concerned with feedback control of AICAR formation and inorganic phosphate/metal transport. *Microbial Cell Factories, 17*(1), 42.

Martin, R. G. (1963). The first enzyme in histidine biosynthesis: the nature of feedback inhibition by histidine. *Journal of Biological Chemistry, 238*(1), 257-268.

Matthews, B. W. (1968). Solvent content of protein crystals. *Journal of molecular biology, 33*(2), 491-497.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., & Read, R. J. (2007). Phaser crystallographic software. *Journal of applied crystallography, 40*(4), 658-674.

Mirabella, A., Terwagne, M., Zygmunt, M. S., Cloeckaert, A., De Bolle, X., & Letesson, J. J. (2013). Brucella melitensis MucR, an orthologue of Sinorhizobium meliloti MucR, is involved in resistance to oxidative, detergent, and saline stresses and cell envelope modifications. *J Bacteriol, 195*(3), 453-65.

Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, Gustavo A., Sonnhammer, E. L. L., Tosatto, S. C. E., Paladin, L., Raj, S., Richardson, L. J., Finn, R. D., & Bateman, A. (2020). Pfam: The protein families database in 2021. *Nucleic Acids Research, 49*(D1), D412-D419.

Mittelstädt, G., Jiao, W., Livingstone, E. K., Moggré, G.-J., Nazmi, A. R., & Parker, E. J. (2018). A dimeric catalytic core relates the short and long forms of ATP-phosphoribosyltransferase. *Biochemical Journal, 475*(1), 247-260.

Mittelstädt, G., Moggré, G. J., Panjikar, S., Nazmi, A. R., & Parker, E. J. (2016). Campylobacter jejuni adenosine triphosphate phosphoribosyltransferase is an active hexamer that is allosterically controlled by the twisting of a regulatory tail. *Protein Science, 25*(8), 1492-1506.

Mittelstädt, G. H. (2015). Allosteric regulation of the adenosine triphosphate phosphoribosyltransferase from campylobacter jejuni.

Moggré, G.-J., Poulin, M. B., Tyler, P. C., Schramm, V. L., & Parker, E. J. (2017). Transition state analysis of adenosine triphosphate phosphoribosyltransferase. *ACS chemical biology, 12*(10), 2662-2670.

Monod, J., Wyman, J., & Changeux, J.-P. (1965). On the nature of allosteric transitions: A plausible model. *Journal of Molecular Biology, 12*(1), 88-118.

Moran, N. A. (2002). Microbial minimalism: genome reduction in bacterial pathogens. *Cell, 108*(5), 583-6.

Morton, D. P., & Parsons, S. M. (1976). Biosynthetic direction substrate kinetics and product inhibition studies on the first enzyme of histidine biosynthesis, adenosine triphosphate phosphoribosyltransferase. *Arch Biochem Biophys, 175*(2), 677-86.

Morton, D. P., & Parsons, S. M. (1977). Inhibition of ATP phosphoribosyltransferase by AMP and ADP in the absence and presence of histidine. *Archives of biochemistry and biophysics, 181*(2), 643-648.

Motlagh, H. N., Wrabl, J. O., Li, J., & Hilser, V. J. (2014). The ensemble nature of allostery. *Nature, 508*(7496), 331-9.

Murshudov, G. N., Vagin, A. A., & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallographica Section D: Biological Crystallography, 53*(3), 240-255.

Notredame, C., Higgins, D. G., & Heringa, J. (2000). T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol, 302*(1), 205-17.

Nussinov, R., Tsai, C.-J., & Ma, B. (2013). The underappreciated role of allostery in the cellular network. *Annual review of biophysics, 42*, 169-189.

Ohta, D., Fujimori, K., Mizutani, M., Nakayama, Y., Kunpaisal-Hashimoto, R., Münzer, S., & Kozaki, A. (2000). Molecular cloning and characterization of ATP-phosphoribosyl transferase from Arabidopsis, a key enzyme in the histidine biosynthetic pathway. *Plant Physiol, 122*(3), 907-14.

Pedreño, S. n., Pisco, J. o. P., Larrouy-Maumus, G. r., Kelly, G., & de Carvalho, L. P. S. r. (2012). Mechanism of feedback allosteric inhibition of ATP phosphoribosyltransferase. *Biochemistry, 51*(40), 8027-8038.

Perutz, M., Rossmann, M., Cullis, A., Muirhead, H., Will, G., & North, A. (1960). Structure of hemoglobin. In *Brookhaven Symp Biol* (Vol. 13, pp. 165-183).

Perutz, M. F. (1970). Stereochemistry of Cooperative Effects in Haemoglobin: Haem–Haem Interaction and the Problem of Allostery. *Nature, 228*(5273), 726-734.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of computational chemistry, 25*(13), 1605-1612.

Pisco, J. P., de Chiara, C., Pacholarz, K. J., Garza-Garcia, A., Ogrodowicz, R. W., Walker, P. A., Barran, P. E., Smerdon, S. J., & de Carvalho, L. P. S. (2017). Uncoupling conformational states from activity in an allosteric enzyme. *Nature Communications, 8*(1), 203.

Prentice, E. J., Hicks, J., Ballerstedt, H., Blank, L. M., Liáng, L. n. L., Schipper, L. A., & Arcus, V. L. (2020). The inflection point hypothesis: The relationship between the temperature dependence of enzyme-catalyzed reaction rates and microbial growth rates. *Biochemistry, 59*(38), 3562-3569.

Ravasio, R., Flatt, S. M., Yan, L., Zamuner, S., Brito, C., & Wyart, M. (2019). Mechanics of Allostery: Contrasting the Induced Fit and Population Shift Scenarios. *Biophysical Journal, 117*(10), 1954-1962.

Remmele, C. W., Xian, Y., Albrecht, M., Faulstich, M., Fraunholz, M., Heinrichs, E., Dittrich, M. T., Müller, T., Reinhardt, R., & Rudel, T. (2014). Transcriptional landscape and essential genes of Neisseria gonorrhoeae. *Nucleic acids research, 42*(16), 10579-10595.

Sause, W. E., Balasubramanian, D., Irnov, I., Copin, R., Sullivan, M. J., Sommerfield, A., Chan, R., Dhabaria, A., Askenazi, M., Ueberheide, B., Shopsin, B., van Bakel, H., & Torres, V. J. (2019). The purine biosynthesis regulator PurR moonlights as a virulence regulator in *Staphylococcus aureus*. *Proceedings of the National Academy of Sciences, 116*(27), 13563-13572.

Schipper, L. A., Hobbs, J. K., Rutledge, S., & Arcus, V. L. (2014). Thermodynamic theory explains the temperature optima of soil microbial processes and high Q10 values at low temperatures. *Global change biology, 20*(11), 3578-3586.

Schwentner, A., Feith, A., Münch, E., Stiefelmaier, J., Lauer, I., Favilli, L., Massner, C., Öhrlein, J., Grund, B., & Hüser, A. (2019). Modular systems metabolic engineering enables balancing of relevant pathways for L-histidine production with Corynebacterium glutamicum. *Biotechnology for biofuels, 12*(1), 65.

Sočan, J., Purg, M., & Åqvist, J. (2020). Computer simulations explain the anomalous temperature optimum in a cold-adapted enzyme. *Nature communications, 11*(1), 1-11.

Stroek, R., Ge, Y., Talbot, P. D., Glok, M. K., Bernaś, K. E., Thomson, C. M., Gould, E. R., Alphey, M. S., Liu, H., Florence, G. J., Naismith, J. H., & da Silva, R. G. (2017). Kinetics and Structure of a Cold-Adapted Hetero-Octameric ATP Phosphoribosyltransferase. *Biochemistry, 56*(5), 793-803.

Studier, F. W. (2005). Protein production by auto-induction in high-density shaking cultures. *Protein expression and purification, 41*(1), 207-234.

Subramanian, S., Jones, H. B., Frustaci, S., Winter, S., van der Kamp, M. W., Arcus, V. L., Pudney, C. R., & Vollmer, F. (2021). Sensing enzyme activation heat capacity at the single-molecule level using gold-nanorod-based optical whispering gallery modes. *ACS Applied Nano Materials, 4*(5), 4576-4583.

Thomas, M. T., & Scopes, K. R. (1998). The effects of temperature on the kinetics and stability of mesophilic and thermophilic 3-phosphoglycerate kinases. *Biochemical Journal, 330*(3), 1087-1095.

Tsai, C. J., del Sol, A., & Nussinov, R. (2008). Allostery: absence of a change in shape does not imply that allostery is not at play. *J Mol Biol, 378*(1), 1-11.

Tzeng, S. R., & Kalodimos, C. G. (2009). Dynamic activation of an allosteric regulatory protein. *Nature, 462*(7271), 368-72.

Van Der Kamp, M. W., Prentice, E. J., Kraakman, K. L., Connolly, M., Mulholland, A. J., & Arcus, V. L. (2018). Dynamical origins of heat capacity changes in enzyme-catalysed reactions. *Nature communications, 9*(1), 1-7.

Vega, M. C., Zou, P., Fernandez, F. J., Murphy, G. E., Sterner, R., Popov, A., & Wilmanns, M. (2005). Regulation of the hetero‐octameric ATP

phosphoribosyl transferase complex from Thermotoga maritima by a tRNA synthetase‐like subunit. *Molecular microbiology, 55*(3), 675-686.

Walton, E. B., & VanVliet, K. J. (2006). Equilibration of experimentally determined protein structures for molecular dynamics simulation. *Physical Review E, 74*(6), 061901.

Weber, G. (1972). Ligand binding and internal equilibiums in proteins. *Biochemistry, 11*(5), 864-878.

Weiner, A. M. (1999). Molecular evolution: Aminoacyl-tRNA synthetases on the loose. *Current biology, 9*(22), R842-R844.

Willenborg, J., & Goethe, R. (2016). Metabolic traits of pathogenic streptococci. *FEBS letters, 590*(21), 3905-3919.

Winkler, M. E., & Ramos-Montañez, S. (2009). Biosynthesis of Histidine. *EcoSal Plus, 3*(2), 10.1128/ecosalplus.3.6.1.9.

Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., & McCoy, A. (2011). Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D: Biological Crystallography, 67*(4), 235-242.

Wolf, Y. I., Aravind, L., Grishin, N. V., & Koonin, E. V. (1999). Evolution of aminoacyl-tRNA synthetases—analysis of unique domain architectures and phylogenetic trees reveals a complex history of horizontal gene transfer events. *Genome research, 9*(8), 689-710.

Wolfenden, R., & Snider, M. J. (2001). The depth of chemical time and the power of enzymes as catalysts. *Accounts of chemical research, 34*(12), 938-945.

Zhang, Y., Shang, X., Deng, A., Chai, X., Lai, S., Zhang, G., & Wen, T. (2012). Genetic and biochemical characterization of Corynebacterium glutamicum ATP phosphoribosyltransferase and its three mutants resistant to feedback inhibition by histidine. *Biochimie, 94*(3), 829-38.

Appendices

Appendix 1: Phylogenetics

A.1. Accession codes (each downloaded sequence)

Table A.1.1. InterPro and UniProt accession numbers used for access to HisG or HisZ amino acid sequences.

Species	Gene	Accession Number
Aquifex aeolicus	HisZ	O67223
Aquifex aeolicus	$HisG_S$	O67543
Bacillus subtilis	HisZ	O34459
Bacillus subtilis	$HisG_S$	O34520
Brucella melitensis	HisZ	P64377
Brucella melitensis	$HisG_S$	Q8YB47
Buchnera aphidicola	$HisG_L$	Q9ZHE7
Caldanaerobacter subterraneus	HisZ	Q8R880
Caldanaerobacter subterraneus	$HisG_S$	Q8R881
Campylobacter jejuni	$HisG_L$	Q5HSJ4
Caulobacter vibrioides	$HisG_L$	B8H5P6
Chlorobaculum tepidum	$HisG_L$	Q8KB10
Clostridium acetobutylicum	HisZ	Q97KI4
Clostridium acetobutylicum	$HisG_S$	Q97KI3
Collinsella aerofaciens	HisZ	A4EBZ1
Collinsella aerofaciens	$HisG_S$	A4EBZ1
Corynebacterium glutamicum	$HisG_L$	Q9Z472
Deinococcus proteolyticus	$HisG_L$	F0RQ13
Deinococcus radiodurans	HisZ	Q9RUE3
Deinococcus radiodurans	$HisG_S$	Q9RUE2
Desulfofarcimen acetoxidans	HisZ	C8W1L4
Desulfuromonas acetoxidans	$HisG_L$	C8VW21
Desulfuromonas acetoxidans	$HisG_S$	C8VW21
Eggerthella lenta	HisZ	C8WPG1
Eggerthella lenta	$HisG_S$	C8WPG1
Escherichia coli	$HisG_L$	P60757
Flavonifractor plautii	HisZ	A0A174I932
Flavonifractor plautii	$HisG_S$	A0A174I955
Geobacter sulfurreducens	HisZ	P60837
Geobacter sulfurreducens	$HisG_L$	P60804
Geobacter sulfurreducens	$HisG_S$	P60836

Haemophilus influenzae	$HisG_L$	P43853
Helicobacter hepaticus	HisZ	Q7VHL1
Helicobacter hepaticus	$HisG_S$	Q7VJU4
Lactococcus lactis	$HisG_S$	Q02129
Lactococcus lactis	HisZ	Q02147
Leptospira interrogans	HisZ	Q8F737
Leptospira interrogans	$HisG_S$	Q8F6P1
Listeria monocytogenes	HisZ	Q722Y1
Listeria monocytogenes	$HisG_S$	C1L0J7
Mesorhizobium japonicum	HisZ	Q987S9
Mesorhizobium japonicum	$HisG_S$	Q987S8
Mycobacterium tuberculosis	$HisG_L$	P9WMN1
Neisseria gonorrhoeae	HisZ	B4RK97
Neisseria gonorrhoeae	$HisG_S$	Q5F7E0
Nostoc punctiforme	HisZ	B2IWE2
Nostoc punctiforme	$HisG_S$	B2IY34
Pasteurella multocida	$HisG_L$	P57919
Pseudomonas aeruginosa	HisZ	B7V202
Pseudomonas aeruginosa	$HisG_S$	B7UZY8
Psychrobacter arcticus	HisZ	Q4FTX3
Psychrobacter arcticus	$HisG_S$	Q4FQF7
Ralstonia solanacearum	HisZ	F6G308
Ralstonia solanacearum	$HisG_S$	F6FXW8
Rhizobium meliloti	HisZ	Q92KL6
Rhizobium meliloti	$HisG_S$	Q92RR6
Salmonella typhimurium	$HisG_L$	P00499
Slackia heliotrinireducens	HisZ	C7N6L4
Slackia heliotrinireducens	$HisG_S$	C7N6L4
Staphylococcus aureus	HisZ	Q2FUT6
Staphylococcus aureus	$HisG_S$	Q2FUT7
Streptococcus sanguinis	HisZ	A3CNT6
Streptococcus sanguinis	$HisG_L$	A3CNT5
Streptomyces coelicolor	$HisG_L$	Q8CK28
Thermosynechococcus elongatus	HisZ	Q8DMD8
Thermosynechococcus elongatus	$HisG_S$	Q8DMD9
Thermotoga maritima	HisZ	Q9X0D3
Thermotoga maritima	$HisG_S$	Q9X0D2
Treponema caldarium	$HisG_L$	F8F494
Vibrio cholerae	$HisG_L$	Q9KSX4
Xanthomonas axonopodis	$HisG_L$	Q8PLH0
Xylella fastidiosa	$HisG_L$	Q9PBC4

Appendix 2: Methods

A.2. HBond script

The Python script used to extract unique H-bonds from ChimeraX H-Bond data (Section 3.10) was kindly developed by Carlin Hamill. Input directory containing H-bond data files to be compared (line 6) and output directory (line 7) must be specified.

```
"""import dependencies"""
import os
import re

"""set working directories"""
input_directory = #Set_input_directory#
output_directory = #Set_output_directory#
log_file = output_directory + '/log.txt'

"""read in data files"""
os.chdir(input_directory)

data_dict = {}#create dictionary for files

for file in os.listdir(input_directory):#iterate through input files
    with open(file) as file_import:#open files
        read_file = [re.split(' +',x) for x in file_import.readlines()][8:]#read files as list of data lines
        data_dict[file] = read_file#add data to dictionary


"""create simplified data for matching"""
match_dict = {}
for file in data_dict.keys():

    hbond_data = data_dict[file]

    temp_list = []

    for idx, entry in enumerate(hbond_data):
        a = entry[0].split('/')[1] #donor chain
        b = entry[1] # donor name
        c = entry[2] # donor residue number
        d = entry[4].split('/')[1] #acceptor chain
        e = entry[5] #acceptor name
        f = entry[6] #acceptor residue number
        temp_list.append([a,b,d,e])

        #dont add number if name is water
        if 'HOH' not in b:
            temp_list[idx].insert(2, c)

        if 'HOH' not in e:
            temp_list[idx].append(f)
```

```
    match_dict[file] = temp_list


for file in data_dict.keys():

    os.chdir(output_directory)

    #generate query file
    query_entries = match_dict[file]

    #appends all other entries to a seperate dictionary for matching
    match_entries = []
    for match_file in data_dict.keys():

        if match_file == file:
            pass
        else:
            for entry in match_dict[match_file]:
                match_entries.append(entry)

    """Perform matching"""

    #searches query against match dictionary and appends if not found
    unmatched = []
    return_query = []
    for idx, query in enumerate(query_entries):

        if query not in match_entries:
            unmatched.append(data_dict[file][idx])
            return_query.append(query)

    """writes log file with all information  on unmatched"""

    #writes output as pymol script to draw bonds
    export = file.split('.')[0] + '_unique.txt'

    with open(log_file, 'a') as log:
        for idx, entry in enumerate(unmatched):
            log.write(f'{idx} |:| {file} |:| {return_query[idx]} |:| {entry}' + '\n')


    """Write outputs"""

    with open(export, 'w') as write_file:

        for idx, entry in enumerate(unmatched):

            #deals with water or his as hbond donor
            if 'hydrogen' in entry:

                acceptor_res = entry[0].split('/')[1]
                donor_res = entry[2]
                donor_atom = entry[3]

                acceptor_chain = entry[4].split('/')[1]
                acceptor_res = entry[6]
                acceptor_atom = entry[7]

                write_file.write(f'distance                    {donor_chain}/{donor_res}/{donor_atom},
{acceptor_chain}/{acceptor_res}/{acceptor_atom}' + '\n')
```

```
        else:

            donor_chain = entry[8].split('/')[1]
            donor_res = entry[10]
            donor_atom = entry[11]

            acceptor_chain = entry[4].split('/')[1]
            acceptor_res = entry[6]
            acceptor_atom = entry[7]



            write_file.write(f'distance                    {donor_chain}/{donor_res}/{donor_atom},
{acceptor_chain}/{acceptor_res}/{acceptor_atom}' + '\n')
```

Appendix 3: Kinetics

A.3.1 IC_{50} L-His and activation by TIH

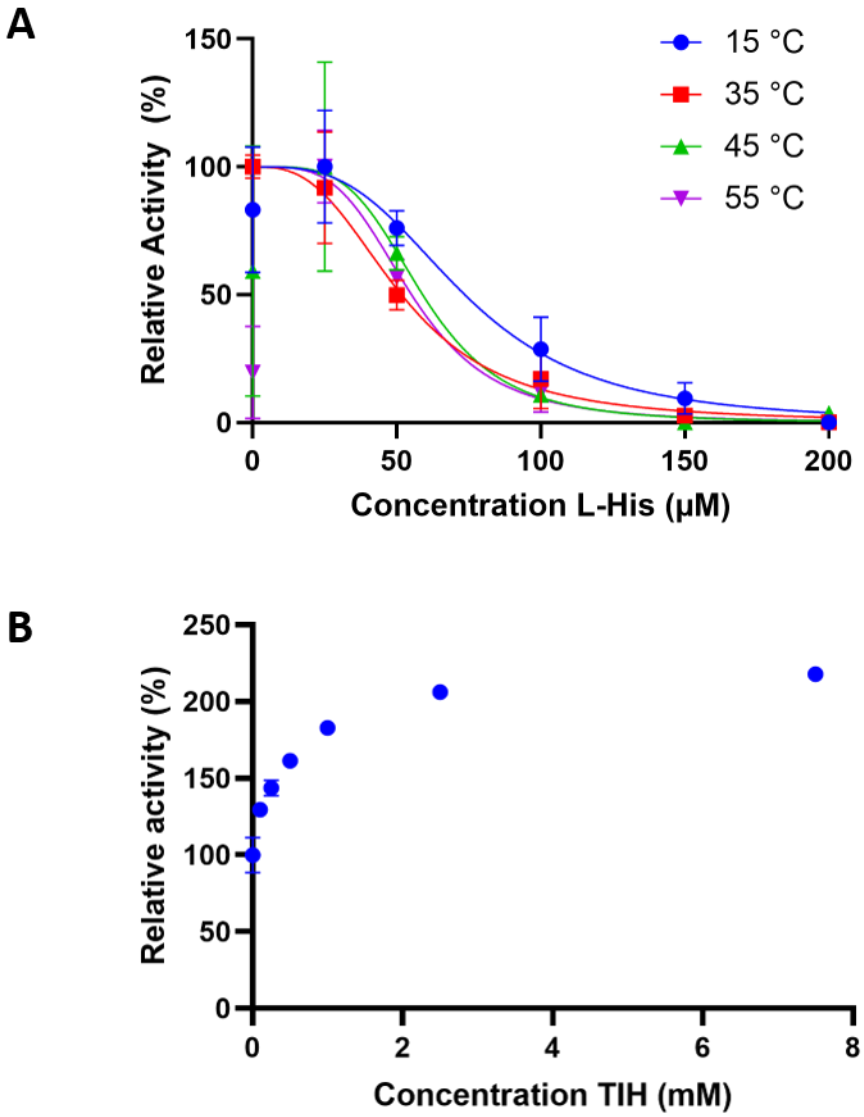

Figure A.3.1 Allosteric ligand response curves of *mtu*ATP-$PRT_{WT.}$ **A.** Normalised data of L-His inhibition across temperature (450 nM *mtu*ATP-PRT, 500 nM *Eco*PPase, 1.8 mM ATP, 1.5 mM PRPP). Mean IC_{50} value of 62 µm. **B.** Dose response kinetic data to TIH activation collected at 37 °C (450 nM *mtu*ATP-PRT, 500 nM *Eco*PPase, 1.8 mM ATP, 1.5 mM PRPP).

A.3.2 Boric Acid Kinetics

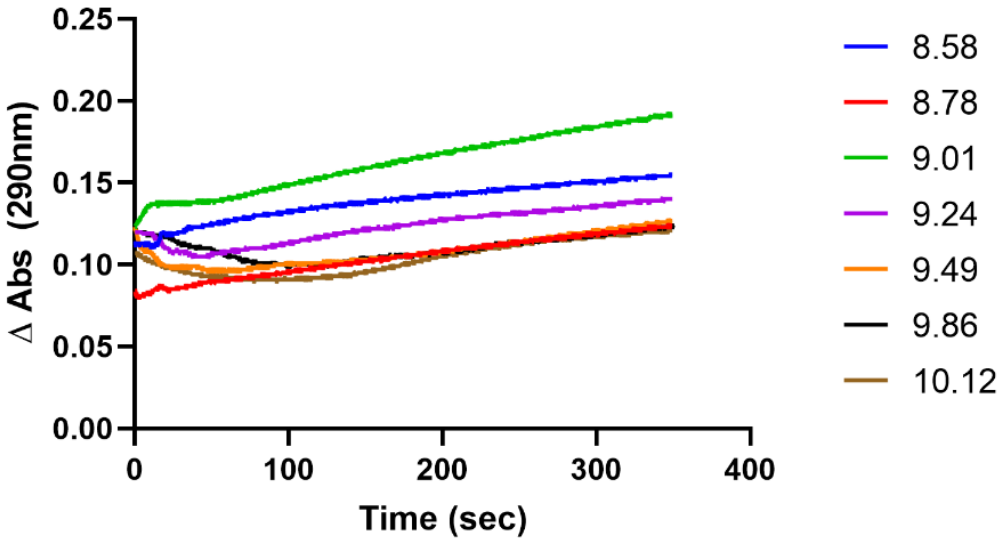

Figure A.3.2. Raw data of *mtu*ATP-PRT$_{WT}$ in boric acid assay buffer (pH variable).

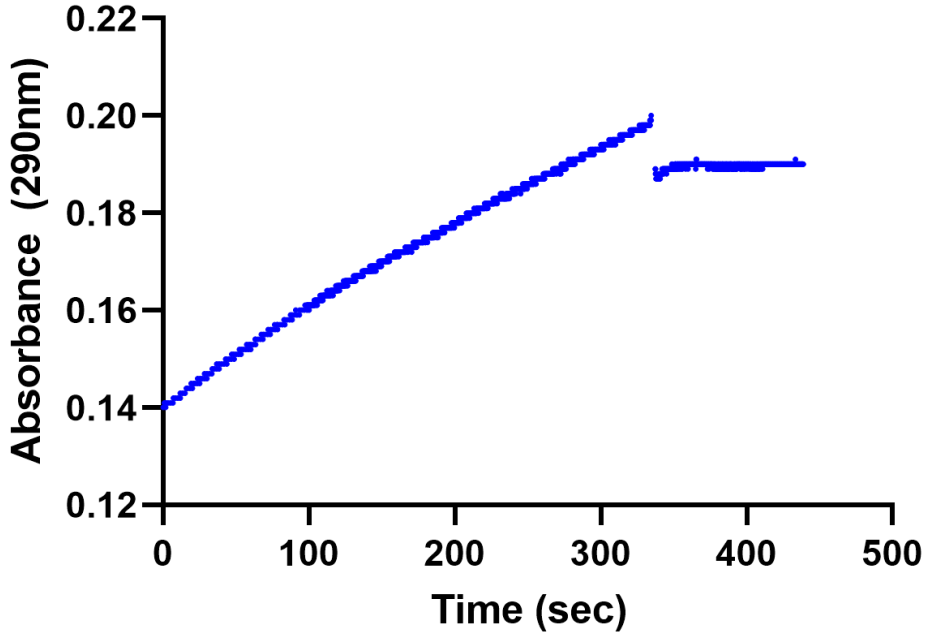

Figure A.3.3. Raw data of *mtu*ATP-PRT$_{WT}$ activity in Boric acid buffer (pH 9.0). Plateau at approximately 350 seconds is in response to the addition of L-His. Plateaued data indicating L-His is a functionally inhibitory ligand in the altered buffering system.

A.3.3 Michaelis-Menten characterisation

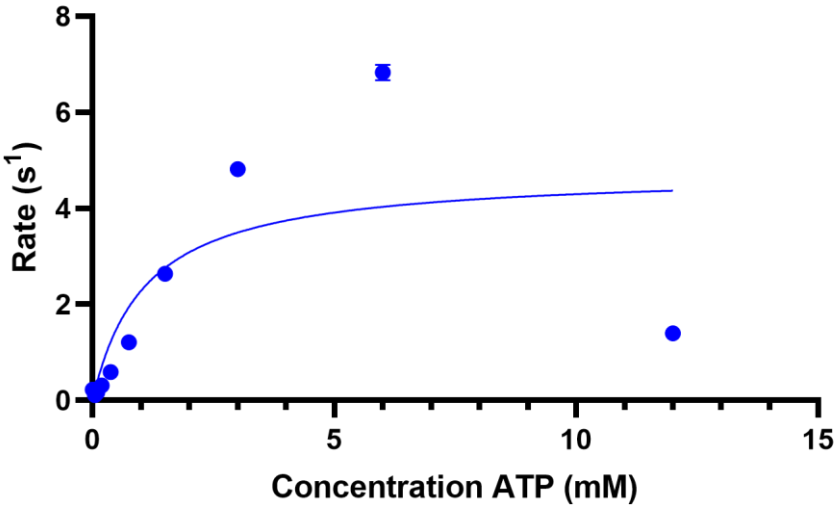

Figure A.3.4. *mtu*ATP-PRT$_{WT}$ Michaelis–Menten plot with varying concentration of ATP. Far greater increased ATP concentrations did not fit typical Michaelis–Menten kinetics, likely due to the increased demand for Mg^{2+} at high concentrations.

A.3.3 MMRT characterisation

Table A.3.1. Temperature independent (unshaded) and temperature dependent (shaded) MMRT fits of *mtu*ATP-PRT variants fit at varying T_0. Given units for; T_0 (K), $\Delta C_p^{\ddagger}$ (J mol $^{-1}$ K $^{-1}$), $\Delta H^{\ddagger}$ (J mol $^{-1}$), $\Delta S^{\ddagger}$ (J mol^{-1}), A (kJ mol^{-1} K^{-2}). Standard Error provided (±).

Mutant	Ligand	T_0	$\Delta C^{\ddagger}_{pT0}$		$\Delta H^{\ddagger}_{T0}$		$\Delta S^{\ddagger}_{T0}$		A		R^2	AICc	T_{opt} (K)	T_{inf} (K)
WT	APO	321	-3900	± 100	16000	± 2000	-177	± 6			0.9918	-204.0	326	310
			-2500	± 400	16000	± 2000	-177	± 5	43	± 11	0.9938	-215.1		
	HIS	323	-3100	± 100	23000	± 1000	-155	± 4			0.9966	-246.3	331	313
			-3600	± 300	22000	± 2000	-158	± 5	-13	± 9	0.9968	-246.1		
	TIH	322	-3600	± 100	11000	± 1000	-188	± 3			0.9974	-270.8	326	309
			-3500	± 200	11000	± 1000	-188	± 3	2	± 7	0.9974	-268.4		
D258A	APO	320	-2900	± 100	15000	± 2000	-189	± 6			0.9878	-211.3	326	308
			-4800	± 300	15000	± 1000	-189	± 4	-63	± 8	0.9947	-248.1		
	HIS	320	-2800	± 100	44000	± 2000	-100	± 5			0.9949	-211.3	337	317
			-5000	± 200	44000	± 1000	-99	± 2	-73	± 5	0.9991	-290.4		
	TIH	314	-4800	± 100	14200	± 800	-184	± 2			0.9917	-239.3	318	304
			-5200	± 100	20000	± 1000	-166	± 4	-39	± 7	0.9952	-263.1		
D258N	APO	323	-3600	± 100	-1000	± 2000	-227	± 5			0.9925	-197.0	323	307
			-5500	± 200	-5000	± 1000	-240	± 3	-49	± 5	0.998	-244.9		
	HIS	320	-3420	± 80	17000	± 1000	-173	± 3			0.9963	-276.8	326	309
			-4200	± 200	18000	± 1000	-169	± 2	-30	± 6	0.9977	-298.1		
	TIH	320	-3300	± 100	19000	± 1000	-162	± 4			0.9947	-239.0	327	309
			-4800	± 200	20000	± 1000	-159	± 3	-50	± 6	0.9981	-285.0		

Appendix 4: Structure

A.4.1 RMSD value changes

Global	WT (APO)	WT (L-His)	WT (TIH)	Δ258A (APO)	Δ258A (L-His)	Δ258A (TIH)	Δ258N (APO)	Δ258N (L-His)	Δ258N (TIH)
WT (APO)	0.00	1.04	1.01	0.45	0.99	1.07	0.36	1.01	1.14
WT (L-His)	1.04	0.00	0.11	1.05	0.43	0.31	0.93	0.27	0.33
WT (TIH)	1.01	0.11	0.00	1.02	0.46	0.33	0.94	0.28	0.32
Δ258A (APO)	0.45	1.05	1.02	0.00	1.10	1.10	0.26	1.10	1.13
Δ258A (L-His)	0.99	0.43	0.46	1.10	0.00	0.46	1.03	0.43	0.45
Δ258A (TIH)	1.07	0.31	0.33	1.10	0.46	0.00	1.02	0.16	0.19
Δ258N (APO)	0.36	0.93	0.94	0.26	1.03	1.02	0.00	1.05	0.93
Δ258N (L-His)	1.01	0.27	0.28	1.10	0.43	0.16	1.05	0.00	0.15
Δ258N (TIH)	1.14	0.33	0.32	1.13	0.45	0.19	0.93	0.15	0.00
Domain I	**WT (APO)**	**WT (L-His)**	**WT (TIH)**	**Δ258A (APO)**	**Δ258A (L-His)**	**Δ258A (TIH)**	**Δ258N (APO)**	**Δ258N (L-His)**	**Δ258N (TIH)**
WT (APO)	0.00	0.62	0.63	0.33	0.61	0.49	0.29	0.51	0.50
WT (L-His)	0.62	0.00	0.08	0.68	0.26	0.29	0.59	0.23	0.24
WT (TIH)	0.63	0.08	0.00	0.70	0.28	0.27	0.58	0.22	0.23
Δ258A (APO)	0.33	0.68	0.70	0.00	0.64	0.46	0.16	0.54	0.48
Δ258A (L-His)	0.61	0.26	0.28	0.64	0.00	0.29	0.62	0.28	0.31
Δ258A (TIH)	0.49	0.29	0.27	0.46	0.29	0.00	0.42	0.22	0.13
Δ258N (APO)	0.29	0.59	0.58	0.16	0.62	0.42	0.00	0.48	0.43
Δ258N (L-His)	0.51	0.23	0.22	0.54	0.28	0.22	0.48	0.00	0.21
Δ258N (TIH)	0.50	0.24	0.23	0.48	0.31	0.13	0.43	0.21	0.00

Domain II	WT (APO)	WT (L-His)	WT (TIH)	Δ258A (APO)	Δ258A (L-His)	Δ258A (TIH)	Δ258N (APO)	Δ258N (L-His)	Δ258N (TIH)
WT (APO)	0.00	0.43	0.46	0.29	0.36	0.51	0.34	0.40	0.53
WT (L-His)	0.43	0.00	0.06	0.48	0.22	0.31	0.35	0.17	0.24
WT (TIH)	0.46	0.06	0.00	0.51	0.21	0.26	0.36	0.18	0.27
Δ258A (APO)	0.29	0.48	0.51	0.00	0.39	0.38	0.19	0.40	0.50
Δ258A (L-His)	0.36	0.22	0.21	0.39	0.00	0.23	0.34	0.25	0.27
Δ258A (TIH)	0.51	0.31	0.26	0.38	0.23	0.00	0.32	0.20	0.12
Δ258N (APO)	0.34	0.35	0.36	0.19	0.34	0.32	0.00	0.37	0.37
Δ258N (L-His)	0.40	0.17	0.18	0.40	0.25	0.20	0.37	0.00	0.23
Δ258N (TIH)	0.53	0.24	0.27	0.50	0.27	0.12	0.37	0.23	0.00

Domain III	WT (APO)	WT (L-His)	WT (TIH)	Δ258A (APO)	Δ258A (L-His)	Δ258A (TIH)	Δ258N (APO)	Δ258N (L-His)	Δ258N (TIH)
WT (APO)	0.00	0.55	0.52	0.23	0.55	0.67	0.26	0.55	0.60
WT (L-His)	0.55	0.00	0.08	0.43	0.27	0.23	0.42	0.20	0.20
WT (TIH)	0.52	0.08	0.00	0.39	0.28	0.23	0.39	0.21	0.22
Δ258A (APO)	0.23	0.43	0.39	0.00	0.47	0.58	0.13	0.47	0.52
Δ258A (L-His)	0.55	0.27	0.28	0.47	0.00	0.27	0.43	0.22	0.25
Δ258A (TIH)	0.67	0.23	0.23	0.58	0.27	0.00	0.55	0.18	0.10
Δ258N (APO)	0.26	0.42	0.39	0.13	0.43	0.55	0.00	0.48	0.52
Δ258N (L-His)	0.55	0.20	0.21	0.47	0.22	0.18	0.48	0.00	0.16
Δ258N (TIH)	0.60	0.20	0.22	0.52	0.25	0.10	0.52	0.16	0.00

Appendix 5: Molecular dynamics

A.5.1 Hbond changes

Table 1.4. Complete hydrogen bond analysis of *mtu*ATP-PRT$_{WT}$ from MD simulations averaged from 3 runs. Occupancy changes greater than 10% are marked in green (bonds formed) and red (bonds broken).

Donor residue	Acceptor residue	WT (APO)	WT (L-His)	WT (TIH)	L-His - APO	TIH - APO	L-His - TIH
GLU61-Main	LEU146-Main	0.0%	46.3%	46.3%	46.3%	46.3%	0.1%
ILE239-Main	PRO241-Main	0.3%	41.3%	37.1%	41.0%	36.9%	4.2%
ARG256-Main	GLN211-Main	39.1%	34.8%	37.5%	-4.3%	-1.6%	-2.7%
ARG280-Main	ALA276-Main	47.7%	24.5%	22.1%	-23.2%	-25.6%	2.4%
ARG251-Side	ASP216-Side	10.6%	22.9%	23.1%	12.2%	12.4%	-0.2%
ARG283-Main	ILE274-Main	33.0%	22.8%	20.6%	-10.2%	-12.3%	2.2%
ARG27-Side	GLU18-Side	0.0%	17.3%	5.0%	17.3%	5.0%	12.3%
SER236-Side	ASP244-Side	1.1%	17.1%	17.2%	16.0%	16.1%	-0.1%
ARG280-Side	GLU22-Side	7.9%	15.3%	14.8%	7.3%	6.9%	0.4%
ARG283-Side	ASP264-Side	24.7%	12.7%	14.8%	-12.0%	-10.0%	-2.1%
ALA276-Main	PHE281-Main	26.3%	12.6%	12.6%	-13.7%	-13.7%	0.0%
ASP278-Main	ASP278-Main	8.5%	11.5%	0.1%	3.0%	-8.4%	11.4%
CYS282-Main	GLY205-Main	14.3%	8.3%	7.6%	-6.0%	-6.8%	0.7%
ARG256-Side	GLN210-Main	9.8%	8.1%	10.4%	-1.8%	0.5%	-2.3%

ARG280-Side	ASP278-Side	6.4%	6.7%	14.3%	0.2%	7.9%	-7.7%
GLN198-Side	ASP264-Side	24.8%	1.2%	3.3%	-23.6%	-21.4%	-2.1%

Table 1.5. Complete hydrogen bond analysis of mtuATP-PRTD258N from MD simulations averaged from 3 runs. Occupancy changes greater than 10% are marked in green (bonds formed) and red (bonds broken).

Donor residue	Acceptor residue	Δ258N (APO)	Δ258N (L-His)	Δ258N (TIH)	L-His - APO	TIH - APO	L-His - TIH
GLU61-Main	LEU146-Main	0.0%	50.4%	50.6%	50.4%	50.6%	-0.2%
ILE239-Main	PRO241-Main	0.8%	17.7%	22.4%	16.9%	21.6%	-4.7%
ARG256-Main	GLN211-Main	40.1%	37.8%	40.5%	-2.3%	0.4%	-2.7%
ARG280-Main	ALA276-Main	33.7%	44.8%	31.7%	11.1%	-2.0%	13.1%
ARG251-Side	ASP216-Side	12.0%	13.9%	19.7%	1.9%	7.7%	-5.8%
ARG283-Main	ILE274-Main	29.0%	30.4%	25.9%	1.4%	-3.2%	4.6%
ARG27-Side	GLU18-Side	0.0%	3.3%	5.3%	3.3%	5.3%	-2.0%
SER236-Side	ASP244-Side	0.9%	6.9%	9.4%	6.0%	8.5%	-2.6%
ARG280-Side	GLU22-Side	12.7%	18.7%	18.9%	6.0%	6.2%	-0.2%
ARG283-Side	ASP264-Side	26.4%	18.5%	20.4%	-8.0%	-6.1%	-1.9%
ALA276-Main	PHE281-Main	44.8%	22.5%	14.1%	-22.3%	-30.8%	8.5%
ASP278-Main	ASP278-Main	0.0%	12.2%	12.1%	12.2%	12.1%	0.1%
CYS282-Main	GLY205-Main	17.5%	7.8%	10.5%	-9.7%	-7.0%	-2.6%
ARG256-Side	GLN210-Main	21.4%	8.4%	7.7%	-12.9%	-13.7%	0.8%
ARG280-Side	ASP278-Side	39.8%	5.8%	7.7%	-34.0%	-32.1%	-1.9%
GLN198-Side	ASP264-Side	22.1%	7.6%	5.4%	-14.5%	-16.7%	2.2%

Table 1.6. Complete comparison of mtuATP-PRTWT and mtuATP-PRTD258N L-His bound structures from MD simulation. Occupancy changes greater than 10% are marked in green (bonds formed) and red (bonds broken).

Donor residue	Acceptor residue	WT (L-His)	Δ258N (L-His)	Δ258N - WT (TIH)
GLU61-Main	LEU146-Main	46.3%	50.4%	4.1%
ILE239-Main	PRO241-Main	41.3%	17.7%	-23.6%
ARG256-Main	GLN211-Main	34.8%	37.8%	3.0%
ARG280-Main	ALA276-Main	24.5%	44.8%	20.4%
ARG251-Side	ASP216-Side	22.9%	13.9%	-9.0%
ARG283-Main	ILE274-Main	22.8%	30.4%	7.6%
ARG27-Side	GLU18-Side	17.3%	3.3%	-14.0%
SER236-Side	ASP244-Side	17.1%	6.9%	-10.2%
ARG280-Side	GLU22-Side	15.3%	18.7%	3.4%
ARG283-Side	ASP264-Side	12.7%	18.5%	5.8%
ALA276-Main	PHE281-Main	12.6%	22.5%	9.9%
ASP278-Main	ASP278-Main	11.5%	12.2%	0.7%
CYS282-Main	GLY205-Main	8.3%	7.8%	-0.5%
ARG256-Side	GLN210-Main	8.1%	8.4%	0.3%
ARG280-Side	ASP278-Side	6.7%	5.8%	-0.9%
GLN198-Side	ASP264-Side	1.2%	7.6%	6.4%

Appendix 6: Isothermal Titration Calorimetry

A.6.1 Isothermal titration calorimetry values

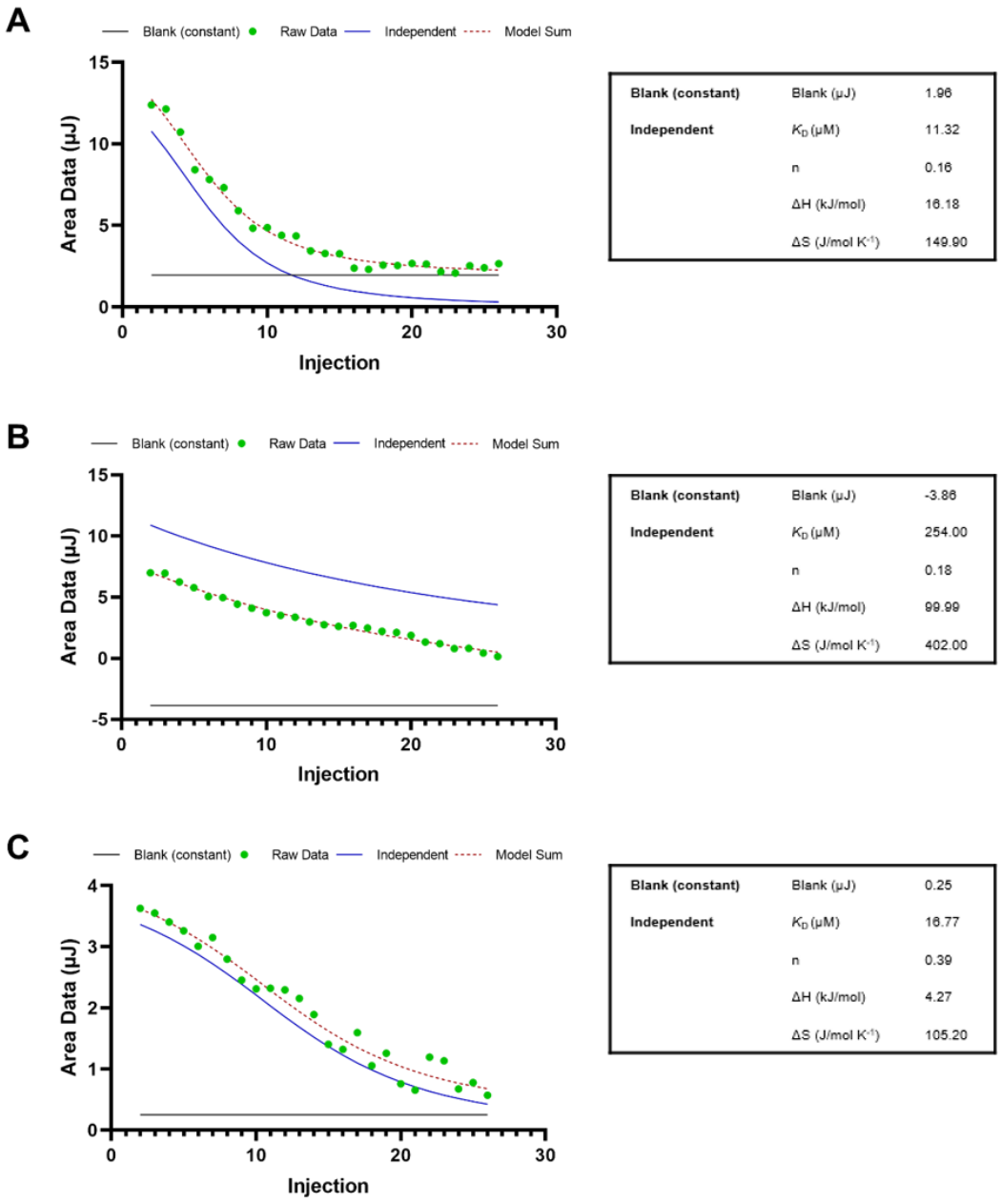

Figure 1.4. Isothermal titration calorimetry of *mtu*ATP-PRT$_{WT}$ (185 µM) titrating up to 1 mM L-His at A. 20 °C B. 27 °C C. 37 °C. Each temperature had a Ki (L-His) of 11.3, 250, 11.6 µM respectively.