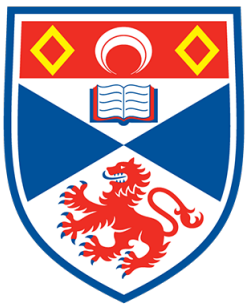
Allostery and Inhibition of *Psychrobacter arcticus* Adenosine 5'-triphosphate Phosphoribosyltransferase.

Catherine Mary Thomson



University of
St Andrews

This thesis is submitted in partial fulfilment for the degree of Doctor of Philosophy (PhD) at the University of St Andrews

May 2019

Declarations

Candidate's declaration

I, Catherine Mary Thomson, do hereby certify that this thesis, submitted for the degree of PhD, which is approximately 29,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for any degree.

I was admitted as a research student at the University of St Andrews in June 2015.

I received funding from an organisation or institution and have acknowledged the funder(s) in the full text of my thesis.

Date Signature of candidate

Supervisor's declaration

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date Signature of supervisor

Permission for publication

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand, unless exempt by an award of an embargo as requested below, that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that this thesis will be electronically accessible for personal or research use and that the library has the right to migrate this thesis into new electronic forms as required to ensure continued access to the thesis.

I, Catherine Mary Thomson, confirm that my thesis does not contain any third-party material that requires copyright clearance.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

Printed copy

Embargo on all of print copy for a period of 1 year on the following ground(s):

- Publication would preclude future publication

Supporting statement for printed embargo request

The data from this thesis is being used to produce a manuscript for future publication.

Electronic copy

Embargo on all of electronic copy for a period of 1 year on the following ground(s):

- Publication would preclude future publication

Supporting statement for electronic embargo request

The data from this thesis are being used to produce a manuscript for furture publication.

Title and Abstract

- I agree to the title and abstract being published.

Date Signature of candidate

Date Signature of supervisor

Underpinning Research Data or Digital Outputs

Candidate's declaration

I, Catherine Mary Thomson, hereby certify that no requirements to deposit original research data or digital outputs apply to this thesis and that, where appropriate, secondary data used have been referenced in the full text of my thesis.

Date Signature of candidate

Acknowledgements

This work was supported by a Wellcome Trust Institutional Strategic Support Fund.

First and foremost, I would like to thank my supervisor, Rafael. G. da Silva for his guidance, support and expertise throughout my PhD, on all three projects. His rapid replies to my many emails and patience over the last four years has been invaluable.

I would never have made it to 'the end' without the unwavering support and unconditional love from my parents, John and Catherine Thomson. They are always quick to remind me how proud of me they are and that I will always have their full approval in any path I choose. Their support, which often came in the form of listening to lengthy rants about the woes of protein purification, was very much appreciated.

I would also like to thank Dr Magnus Alphey for the many hours he dedicated to showing me the 'crystallography-ropes' and never once complaining about how chilly it was in the 4°C cold room while I tried, and tried, to fish out my first crystal. A further thank you must go to current and past members of the da Silva lab. In particular, thanks go to Gemma Fisher whose work is closely aligned to mine and often, during times of purification troubleshooting, became nearly one and the same.

My appreciation also goes to Dr Clarissa Czekster for the many times she helped me in the lab, but in particular for my second project, which unfortunately remains on the cutting room floor. I would also like to extend my gratitude to Fiona McKissock and Ying Ge, both of whom convinced me to convert my thesis to LaTeX. To Ying for getting me started with my preamble and to Fiona for helping me figure out it's 'quirks', which was often during the early hours of the morning.

To Matthew, Anna, Amy and my closest friends: I thank you for your encouragement, patient ears and the good times we shared while it often felt like I would never achieve a positive lab result.

Finally, thanks go to Ashleigh for routinely reminding me that light can be found in the darkest of places, if one only remembers to turn on the light.

Abstract

Allostery is the regulation of a protein by binding in an area distinct from the enzyme active site, but despite having clear applications in biotechnology and in drug discovery, elucidation of it's complexities are still ongoing. Allostery can be either activating or inhibiting and the allosteric site is often distant to the one affected. The enzyme *Psychrobacter arcticus* ATP phosphoribosyl transferase (ATPPRT) is both allosterically activated and inhibited and is therefore useful as a model organism for understanding allostery. *P. arcticus*, ATPPRT is a Type IV ATPPRT, catalyzes the first step in histidine biosynthesis, and is comprised of two subunits: a short form $HisG_S$ and HisZ. $HisG_S$ is the catalytic domain and is allosterically activated by HisZ. HisZ, also contains an allosteric binding site for inhibition by histidine. Short form $HisG_S$ alone is catalytically active but insensitive to histidine. This work establishes the kinetic mechanism of histidine inhibition in *Pa*ATPPRT, with histidine binding non-competitively with respect to both substrates (ATP and PRPP). AMP was confirmed as an competitive inhibitor with respect to both substrates for both *Pa*ATPPRT and *Pa*$HisG_S$ and ADP was determined to be an alternative substrate. This work demonstrates the first example of histidine binding in a site analogous to that of HisRS in a short-form ATPPRT. It also highlights an interaction between two loops in *Pa*HisZ, mediated by a hydrogen bond between Tyr263 and His104, that is likely to greatly reduce the probability of the complex sampling the activated conformation necessary for successful reaction. Additionally, a clear distinction between the activated form of *Pa*ATPPRT with ATP bound compared to all other structures is observed. Overall, this work demonstrates a clear reliance on key structural changes for the allosteric regulation of *Pa*ATPPRT.

Attempts to purify *Staphylococcus aureus* ATPPRT were unsuccessful.

Publications

- **Thomson, C.M.**, Alphey, M. S., Fisher, G., and da Silva, R. G. (2019) Mapping the structural path for allosteric inhibition of a short-form ATP phosphoribosyltransferase by histidine. *Biochemistry* 58, 3078-3086.
- Fisher, G., **Thomson, C.M.**, Stroek, R., Czekster, C. M., Hirschi, J. S. & Silva, R. G. (2018) Allosteric Activation Shifts the Rate-Limiting Step in a Short-Form ATP Phosphoribosyltransferase. *Biochemistry* 57, 4357–4367.
- Stroek, R., Ge, Y., Talbot, P.D., Glok, M.K., Bernas, K.E., **Thomson, C.M.**, Gould, E.R., Alphey, M.S., Liu, H., Florence, G. J., *et al.* (2017) Kinetics and Structure of a Cold-Adapted Hetero-Octameric ATP Phosphoribosyltransferase. *Biochemistry* 56, 793-803.

Contents

Abbreviations

6xHis-tag	six residue histidine tag
A. baumannii	*Acinetobacter baumannii*
ADP	adenosine 5'-diphosphate
AICAR	5-aminoimidazole-4-carboxamide ribonucleotide
AMP	adenosine 5'-monophosphate
ATP	adenosine 5'-triphosphate
Auto ind.	auto induction
B. thailandensis	*Burkholderia thailandensis*
Bis-Tris	2,2-Bis(hydroxymethyl)-2,2',2" nitrilotriethano
BMCE	beta-mercaptoethanol
C. jejuni	*Campylobacter jejuni*
CPS	cyber physical system
CTP	cytidine triphosphate
CV	column volume
Da	Dalton
dATP	deoxyadenosine 5'-triphosphate
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTPs	deoxynucleotide mix
DSF	differential scanning fluorimetry
DTT	dithiothreitol
E	enzyme
E. coli	*Escherichia coli*
EC	enzyme commission
EDTA	ethylenediaminetetraacetic acid
ESI-MS	electrospray ionisation mass spectrometry
FPLC	fast protein liquid chromatography
FU	fraction unfolded
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HF	high fidelity (polymerase/buffer)
HisG$_L$	long form HisG
HisG$_S$	short form HisG

I	inhibitor
IC_{50}	half-maximal inhibitory concentration
IPTG	isopropyl β-D-1-thiogalactopyranoside
ITP	inosine 5'-triphosphate
k_{cat}	turnover number
K_D	dissociation constant
K_i	inhibitor constant
K_M	Michaelis constant
L. lactis	*Lactococcus lactis*
LB	lysogeny broth
M. tuberculosis	*Mycobacterium tuberculosis*
MBP	maltose-binding protein
***Mt*PPase**	*Mycobacterium tuberculosis* pyrophosphatase
MWCO	molecular weight cut off
NCBI	National Centre for Biotechnology Information
OD	optical density
P	product
P. arcticus	*Psychrobacter arcticus*
***Pa*ATPPRT**	*Psychrobacter arcticus* ATP phosphoribosyl transferase
***Pa*HisG$_S$**	*Psychrobacter arcticus* His Gs (short form)
***Pa*HisZ**	*Psychrobacter arcticus* His Z
PCR	polymerase chain reaction
PPase	pyrophosphatase
PRATP	N^1- (5-phospho-β-D-ribosyl)-ATP
PRPP	5-phospho-α-D-phosphoribosyl-1-pyrophosphate
r.m.s.d	root mean square deviation
R5P	ribose-5-phosphate
S	substrate
S. aureus	*Staphylococcus aureus*
S. typhimurium	*Salmonella typhimurium*
***Sa*ATPPRT**	*Staphylococcus aureus* ATPPRT phosphoribosyl transferase
***Sa*HisG$_S$**	*Staphylococcus aureus* His G_S (short form)
***Sa*HisZ**	*Staphylococcus aureus* His Z
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SOC	super optimal broth
T	temperature
T. maritima	*Thermotoga maritima*
TBE	Tris/borate/EDTA buffer
TEVP	tobacco etch virus protease
TIH	3-(2-thienyl)-L-alanine
T_m	melting temperature
U	unit
UV	ultraviolet
v	kinetic rate
WT	wild type

List of Figures

List of Tables

Chapter 1

INTRODUCTION

1.1 Allostery

Allostery, although regarded as "the second secret of life"[1, 2], is still not completely understood. Despite having clear applications in biotechnology [3] and in drug discovery [4], elucidation of it's complexities are still ongoing. There is a consensus over what allostery is and what allostery can do, yet several details of its intricacies remain elusive.

1.1.1 Allostery: The Basics

Allostery is the regulation of a protein by binding in an area distinct from the active site [5]. For example, Figure 1.1 shows a simple allosteric inhibitor. The allosteric inhibitor binds to a second binding site (not the active site), which in this case, causes a conformational change that prevents substrate binding. In theory, this means that aside from the active site, any area inside or on the surface of the protein is a potential allosteric effector binding site [6].

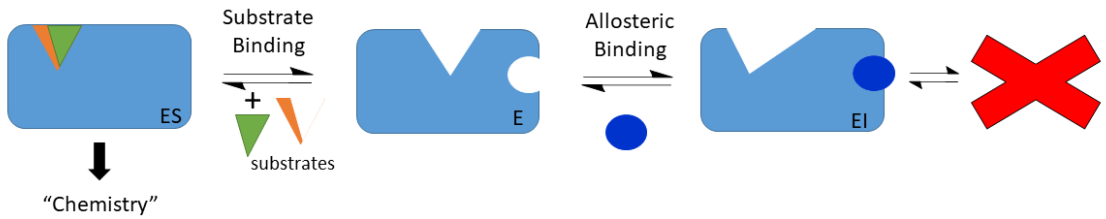


Figure 1.1: Schematic representation of an allosteric inhibitor binding in a secondary site, distorting the active site and preventing substrate binding.

Sometimes allostery and cooperativity are used interchangeably but they are in fact

distinct concepts, though related. Cooperativity is concerned with regulation of a multimeric protein through change in its quaternary structure from one subunit to the same site on another subunit [5], and therefore involves allostery. Historically, cooperativity was also referred to as allostery, however, allostery can involve any site (except the active site) that causes a regulatory effect elsewhere in the same molecule and thus allostery is possible in monomeric proteins too. In other words, all cooperativity is allosteric but not all allostery is cooperative. Allostery can also be ligand-free, such as regulation through amino acid mutations [7] or post-translational modification [8].

Cooperativity can be either positive, where binding of a ligand to one subunit *increases* the affinity of binding to another subunit (Figure 1.2, A) or negative, where binding to one subunit *decreases* the affinity of binding to another subunit (Figure 1.2, B)[5].

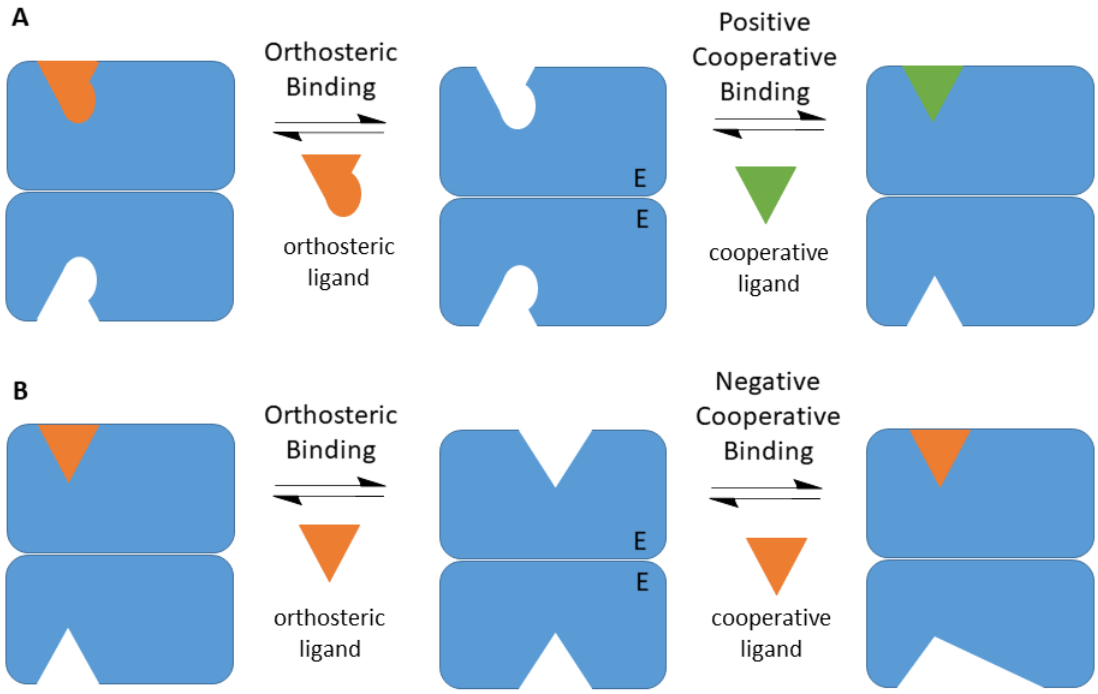


Figure 1.2: Schematic representation of positive (A) and negative (B) cooperativity. As the positive cooperativity ligand binds, it induces a conformational change in the neighbouring active site that increases affinity for subsequent ligand binding. As the negative cooperativity ligand binds, it induces a conformational change in the neighbouring active site that decreases affinity for subsequent ligand binding.

Examples of molecules displaying positive cooperativity are oxygen towards haemoglobin [9], mannose towards glucokinase [10] and estradiol towards the estrogen receptor [11]. An examples of a protein with negative cooperativity is CTP synthase and its substrate glutamine [12].

Allostery too, can produce positive or negative modulation. Allosteric activation occurs

when binding of an allosteric ligand to any non-active site in the monomeric or multimeric protein increases the affinity and/or activity with substrate or other ligand in their respective active sites. Examples of allosteric activators include ribosomal S6 kinase 2 (Rsk2) to the hormone-binding domain of estrogen receptor alpha [13] and heterodimer formation of *Trypanosoma brucei* S-adenosylmethionine decarboxylase with a catalytically dead homolog [14]. Conversely, allosteric inhibition occurs when binding of an allosteric ligand to any non-active site in the mono- or multimeric protein decreases the affinity and/or activity with substrate or other ligand in their respective active sites. Examples of allosteric inhibitors include phosphoenolpyruvate with respect to *Escherichia coli* (*E. coli*) phosphofructokinase [15] and CTP with respect to aspartate transcarbamoylase [16–18]. Some enzymes can be both allosterically activated and inhibited, for example, synthases such as *Arabidopsis* threonine synthase [19] and ATP phosphoribosyltransferases (ATPPRT) such as *Psychrobacter arcticus* ATPPRT (*Pa*ATPPRT) [20].

1.1.2 Applications of Allosteric Modulation

The ability to modulate activity without interfering directly with the active site creates broad scope for academic and industrial applications, particularly in biotechnology and drug discovery.

Allostery in Biotechnology

With a global shift towards sustainability and environmentally robust practises, biology-based methodologies are becoming extremely valuable. From enhancing the quality and quantity of crops [21], to utilising low temperatures and organic-solvent free systems for chemical synthesis [22], biotechnology is a fast growing and lucrative sector. For example, over 4 million tons of L-amino acids are bought globally per year [23]. Many of these amino acids, such as L-glutamate [24] and L-histidine [25] are still produced by laborious purification from bacterial strains.

There is a clear economic and environmental advantage to engineering enzymes and organisms to produce desirable chemicals faster and more efficiently. Given that many enzymes are allosterically modulated, for example by feedback inhibition [16, 26], understanding allosteric mechanisms is of great importance to overcome and engineer allostery as needed.

Allostery in Drug Discovery

Allostery is not only of use to the synthetic biology and biotechnology industries, it is also of interest to the pharmaceutical industry. Allosteric drugs, have been gaining traction over the last 5-10 years, a trend which can be visualized by analyzing the number of papers mentioning "allosteric drug" from 1997 to 2017 (Figure 1.3). There is a distinct rise from 4 papers to over 300 in only two decades with numerous allosteric drug candidates now in clinical trials [27–30].

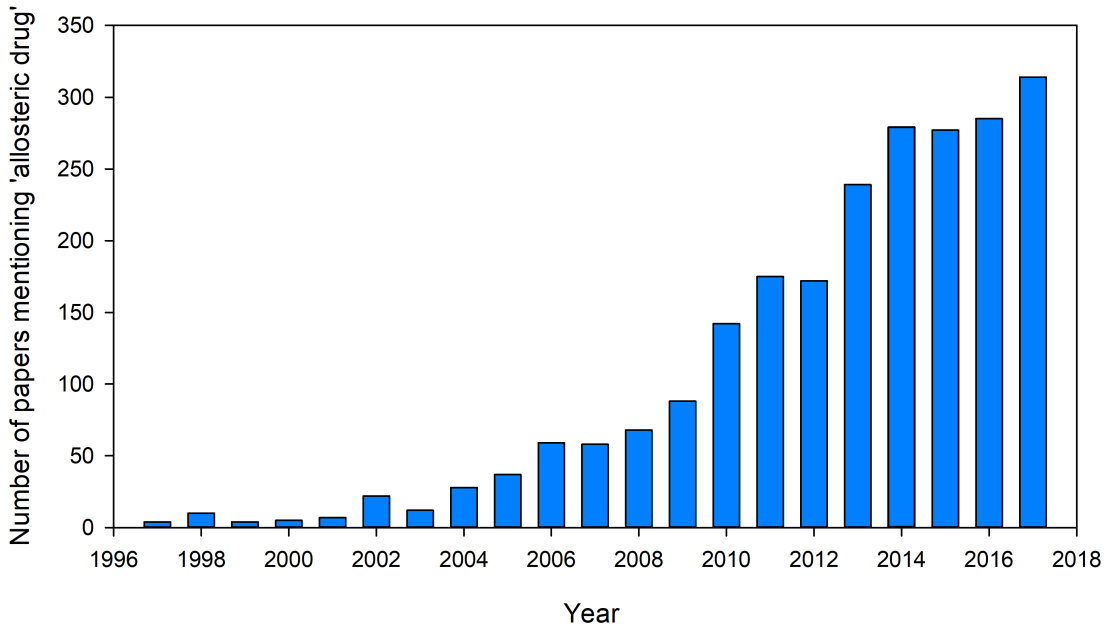


Figure 1.3: Number of publications per year that mention "allosteric drug" [31]

Although rationally designed allosteric drugs are slowly coming into the clinic, it is orthosteric drugs that have historically dominated the pharmaceutical industry. Orthosteric drugs bind directly to the active site of the target and directly modulate activity by blocking further binding at that site and inhibiting the target's function. This makes orthosteric drug targets relatively easy to assay given that the ligand interacts directly with the active site and the corresponding enzyme function. However, orthosteric drugs generally result in complete inhibition or activation of the target rather than a tunable modulation. They are also more prone to adverse side effects in patients as many targets have similar binding sites to other enzymes with related functions [32]. Allosteric drugs on the other hand, have a better chance of targeting subtype specificity and reduce off-target effects as their binding sites are less conserved than their orthosteric counterparts. Additionally, allosteric drugs have increased flexibility to target a site that is only open to binding during specific times in activation or signalling [33]. Unfortunately, the current limitation for allosteric drug discovery is mainly due

to current drug design techniques being poorly suited to reporting weaker-binding allosteric modulators [33]. However, there are some techniques, such as fragment-based drug design, which are more suited for weaker-binding allosteric modulators.

However, allosteric drugs have actually been on the market since the early 1900s, despite their mode of action being unknown at the time of introduction. For example, the barbiturate, diethyl-barbituric acid, was first prescribed [34] in 1904. It has since been determined that barbiturates (such as phenobarbital, still prescribed as an anti-convulsant) allosterically stimulate the inhibitory neurotransmitter γ-aminobutyric acid (GABA) and cause an increase in the length of time the GABA channel is open [35]. Similarly, benzodiazepines introduced in the 1960s, such as Diazepam, also allosterically regulate GABA but by increasing the frequency of GABA channel opening [36] to relieve anxiety in patients. Interestingly, both barbiturates and benzodiazepines had been in clinical use routinely while allostery, as a concept, was just in its infancy.

1.1.3 Brief Historical Perspective of Allostery and Cooperativity

The term of "allostery" was only coined around 50 years ago. In the 1950s, covalent modification of enzyme activity and feedback inhibition of enzymes were discovered [37], and in 1961 'allosteric sites' was used by Jacques Monod and Francois Jacob when describing the end-product inhibition of an enzyme where the inhibitor was not a structural analogue of the substrate [38]. A few years later, in 1965, the first model of cooperativity was put forward by Monod, Wyman, and Changeux and is referred to as either the symmetry model or the MWC model [39]. However, Pauling had described a similar phenomenon in 1935 [40] when proposing a model relating to the "Bohr effect", the term used ca. 1903 to explain how higher concentrations of carbon dioxide could inversely affect oxygen's binding affinity for haemoglobin [39]. An alternative model was proposed by Koshland, Nemethy and Filmer [41], referred to as the sequential model or KNF model. Both the symmetry and sequential models are limiting cases of cooperativity. They were conceived to explain haemoglobin's binding mechanism, and in many cases, cooperativity is unlikely to be a clear-cut example of either model. These models are still useful and used in active research [42], but with the caveat that lots of intermediate cases exist. It is also important to note that in the early decades of allostery research, 'cooperativity' was used interchangeably with 'allostery'. The former is now regarded as only one type of allostery.

Symmetry Model (MWC model)

The symmetry model (Figure 1.4) of positive cooperativity acknowledges that a symmetrical oligomeric protein exists in equilibrium between one of two stable forms

of free enzyme, tensed state (T) and relaxed state (R) [43]. Ligand binding to one form displaces the equilibrium towards that form. However, the symmetry model maintains that the multimeric protein is a symmetrical oligomer and therefore subunits in the same oligomer have the same conformational state. All subunits will be either in the R state or in the T state. There can be no mixture of T and R state within one protein oligomer [44]. Therefore, the symmetry model describes concerted allosteric transitions of proteins to maintain identical subunits within a given oligomer. These allosteric transitions occur between the free protein oligomers, are independent of ligand-binding and can only account for positive cooperativity.

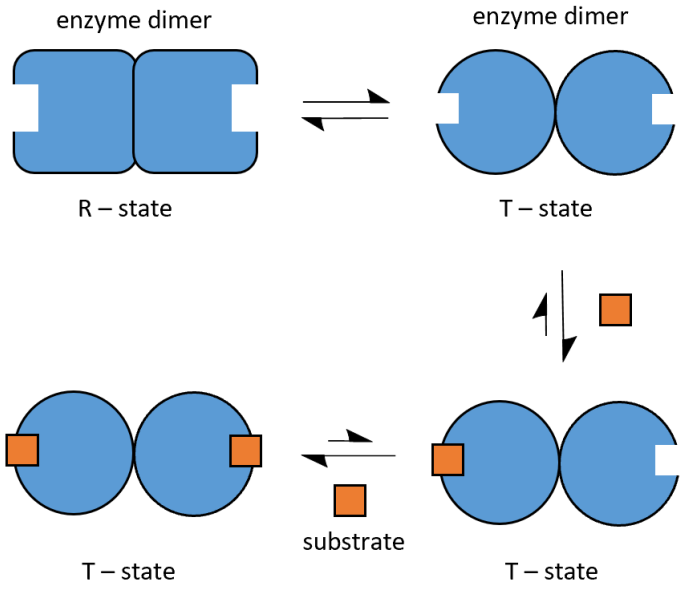


Figure 1.4: Schematic representation of the symmetry model of cooperativity. Enzyme as a whole exists in two states, T and R, and can change freely between them. Ligand binding displaces the equilibrium towards one state. In this example binding is preferred for the T- state (small square binding site) over the R-state (large square binding site). Since the T-state is favoured, ligand binding by a second ligand is more favourable after conformational rearrangement to the T state.

Sequential Model (KNF model)

The sequential model (also referred to as the induced fit model or the KNF model), (Figure 1.5) contrary to the symmetry model, postulates that there is only one stable form of free enzyme in solution. It states that when a cooperative ligand binds to one subunit of the multimeric protein, it induces a conformational change in the ligand-bound subunit [41]. This conformational change is then able to affect a neighbouring subunit, or subunits, making it more or less likely for a second ligand to bind in neighbouring binding sites. This creates an asymmetrical oligomer capable of having subunits in more than one state within the same oligomer [44]. The changing of one subunit to a high

affinity state, for example, makes neighbouring subunits more likely to transition to a higher affinity state due to a lowering of the energy barrier correlating to that state transition [44]. The sequential model also holds in reverse for negative cooperativity: binding of a molecule to one subunit causes a conformational change from high to low binding affinity in the other subunits. This is contrary to the symmetry model which can only account for positive cooperativity.

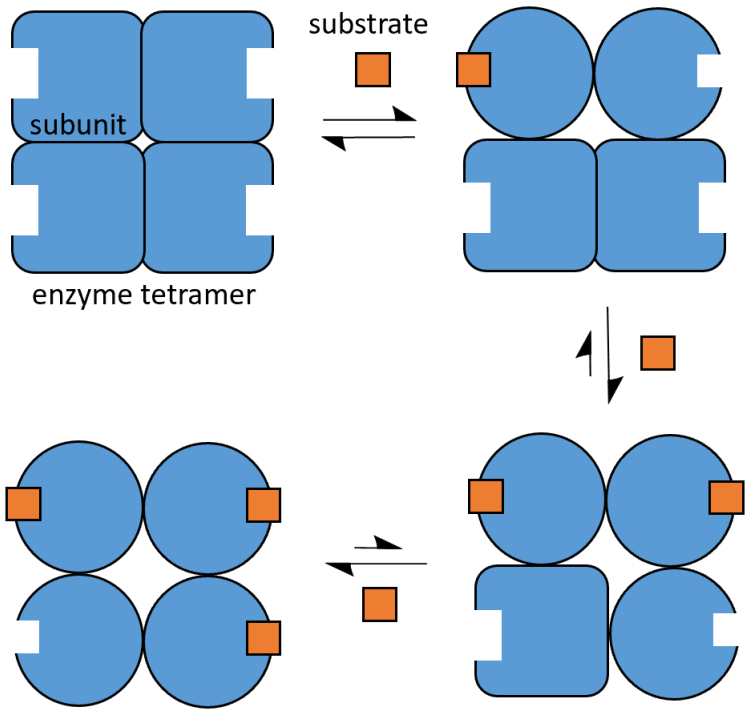


Figure 1.5: Schematic representation of the sequential model of cooperativity. Each enzyme subunit can exist in two interchangeable states. Ligand binding causes a conformation change in the neighbouring subunit. In this example, binding is preferred for the circle state (with small square binding site) over the square state (large square binding site). Since the circle state is favoured, ligand binding by a second ligand is more favourable after conformational rearrangement to the circle state.

Contemporary Views on Allostery

For almost 20 years the symmetry and sequential models were the best explanation for cooperativity (with allostery as a whole being somewhat overlooked). They still have merit as extreme cases, but the mechanism of cooperativity and allostery without cooperativity has moved forward. The development of X-ray structures with and without ligands [45] lent support to structural changes occurring upon ligand binding whereas Cooper and Dryden [46] proposed that allosteric regulation was also possible without any major conformational change, but using a shift in protein dynamics instead [47]. It

is likely that both are possible but structural changes are the significantly more common mechanism for allosteric regulation due to inherent flexibility, of at least some regions, in almost all proteins [48, 49].

Views on allostery with and without cooperativity have expanded to acknowledge the ensemble nature of proteins [50]. Protein folding is not always an all-or-none approach between two states. For small proteins this can be the case but for large proteins they can exist in varying degrees of unfolding or intermediate structures between their main conformations [51]. These conformational states, or structural ensemble, have a high degree of conformational freedom and their energy landscape has many local energy minima [52]. Consequently, there are many possible structural conformations available for ligand-binding to occur. With this view, allostery can be considered as a process whereby the flexibility in proteins creates populations of conformers that can interconvert. The binding of a molecule initiates a redistribution of protein conformational ensembles capable of altering the proteins interconversion possibilities [8]. Kern et.al. [53] have elegantly shown through NMR and binding kinetics that enzyme affinity between kinases can be a direct result of differences in conformational equilibrium post-binding, rather than differences in the binding step specifically. Gleevec, the first kinase inhibitor to reach the market for cancer [54], is a potent drug used to treat chronic myeloid leukaemia. Puzzlingly, despite almost identical binding sites in the tyrosine kinase Abl subfamily and the closely related Src kinase subfamily, Gleevec has an affinity 3000 times stronger for Abl [55]. The difference in affinity is not due to amino acid differences in the drug binding pocket, or to a steric clash in a highly conserved loop of Src, absent in Ab1 [55]. It was instead shown that the drastic difference in binding affinities was due to conformational selection before binding and an induced fit step after binding. Gleevec binds to a conformation sampled in solution much more frequently by Ab1 than Src. Subsequent to the conformational change, there is no noticeable difference between the two crystal structures as the binding site is identical after the induced fit occurs and under crystallographic conditions [53]. Thus, it is the significant equilibrium difference stemming from the ensemble nature of proteins that distinguishes Abl from Src in Gleevec binding. The consequences of this experiment, and others [56, 57], is an understanding that the simplistic symmetry and sequential models that assume proteins in their native state are only in equilibrium between stable and discrete conformations does not represent the full picture. It is now widely accepted that proteins are continuously sampling among multiple conformations across a complex energy landscape, with unfolding and refolding simultaneously occurring in localized areas [8]. These redistributions include those brought into fruition by allosteric effector binding and it is common in proteins that some states will exchange very quickly but other conformations will take much longer to be sampled because the barrier for the transition is much higher. However, the limiting cases of the symmetry and sequential models for cooperative binding remain relevant to research with an

appreciation that the specific states pertinent for allostery with cooperativity within the symmetry and sequential models are likely to correspond to the favourable states in the energy landscape.

1.1.4 Feedback Inhibition

The ensemble nature of proteins is also relevant for the explanation of feedback inhibition, another type of allostery not necessarily involving cooperativity. Feedback inhibition, or end-product inhibition, is the negative control of a molecule's synthesis by the last product in the pathway inhibiting the first committed step in the biosynthesis (Figure 1.6) [5]. Feedback inhibition is a nearly instantaneous effect that is independent of transcription, translation and signal transduction cascades, allowing it to be one of the most efficient flux control mechanisms in biochemical pathways[58–60]. This makes study of feedback inhibitors useful for advancement in allosteric small molecule design, chemical and genetic tools, synthetic biology regulation and disease therapeutics[58].

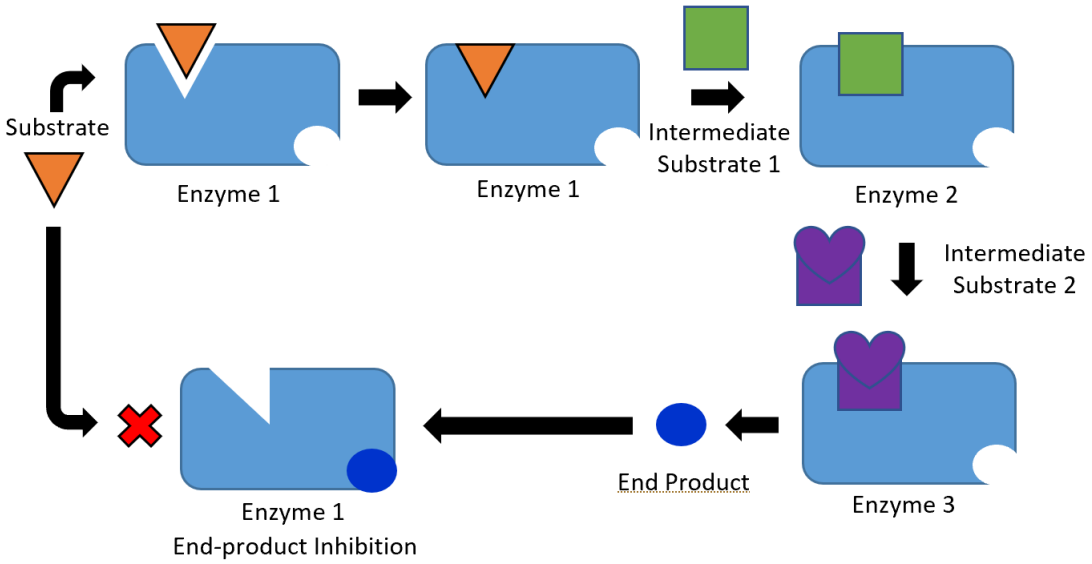


Figure 1.6: Schematic representation of end-product inhibition in a three-step enzyme pathway. Substrate 1 (orange triangle) binds to enzyme 1 to form intermediate substrate 1 (green square). Intermediate substrate 1 can bind to enzyme 2 to form intermediate substrate 2 (purple shape). Intermediate substrate 2 can bind to enzyme 3 to form the end product (blue circle). The end product is capable of inhibiting enzyme 1.

Feedback inhibition allows the pathway to regulate and control the synthesis of product formation, so overproduction does not occur. When product is present in the cell (and therefore not being used up immediately) feedback inhibition signals a cessation or reduction of the molecule's production by binding to the initial enzyme in the pathway and inhibiting or slowing down further synthesis [5]. Reversible binding allows the

production to begin again at a time when the levels of the end-product decrease. This saves valuable cell resources such as carbon, nitrogen, and energy from unnecessary assembly of products not immediately required [60]. Since feedback inhibition directly utilises product to affect the starter enzyme it may be supposed that orthosteric inhibition would be the simplest way for product binding, with substrate and product vying for the same binding site. However, with pathways many steps long, the product is often very chemically distinct from its initial substrate [60]. This poses a challenge for orthosteric inhibition since accommodating both substrate and product in high affinity, across the diverse array of flux-controlling enzymes, is not trivial. Allostery is an effective tool to overcome this. Interestingly, feedback inhibition was discovered before the concept of allostery was formulated and before the focus of its explanation by cooperativity made the forefront of the field. The first instance of feedback inhibition was reported in 1954 by Brooke, Ushiba and Magasanik regarding feedback inhibition in bacterial pyrimidine biosynthesis [61] with supporting evidence following closely in the year after. For example, Novick and Szilard showed that the first enzyme in the system was responsible for this control [62]. Similarly, Roberts *et al.* conducted a study using *E. coli* with all carbon compounds in the cell ^{14}C-labelled [63]. Upon addition of unlabelled end-product, endogenous biosynthesis was down-regulated with the cell exclusively favouring use of the exogenous end-product source. Following these key observations, many systems were interrogated and found to have feedback inhibition. The most studied of these, aspartate carbamoyltransferase (ATCase), was demonstrated to control pyrimidine biosynthesis by measuring activity under exogenously controlled end-product supply [64]. Feedback inhibition is a quick and simple way to control metabolism and has since been found to be widespread. Feedback inhibition is abundant in bacterial biosynthetic pathways for products such as amino acids, nucleotides and even isoprenoids [65]. Almost all bacterial amino acid biosynthesis utilises feedback inhibition. For example, tryptophan biosynthesis [66], valine biosynthesis [67] and histidine biosynthesis [68].

1.2 Histidine Biosynthesis and the Role of ATP Phosphoribosyltransferase

1.2.1 Histidine Biosynthesis Pathway

The histidine pathway is a nine-step biosynthesis (Figure 1.7) beginning with the Mg^{2+}-dependent nucleophilic attack of adenosine 5'-triphosphate (ATP) N1 on 5-phospho-α-D-phosphoribosyl-1-pyrophosphate (PRPP) C1 catalysed by the flux-controlling enzyme ATP phosphoribosyltransferase (ATPPRT) (EC: 2.4.2.17) (Figure 1.8). This reversible reaction producing N^1- (5-phospho-β-D-ribosyl)-ATP (PRATP) and inorganic pyrophosphate (PP_i) strongly favours reactants but is followed by eight further steps to synthesise L-histidine. Hydrolysis of the α – β phosphate linkage, then further hydrolysis to open the six membered ring to give phosphoribosyl-AMP, is carried out by the bifunctional phosphoribosyl-ATP pyrophosphohydrolase/phosphoribosyl-AMP cyclohydrolase (EC: 3.6.1.31/EC: 3.5.4.19) in some organisms (e. g. *Psychrobacter arcticus* and *Staphylococcus aureus*), while in others (e. g. *Mycobaterium tuberculosis* and *Escherichia coli*) each activity is present in a unique polypeptide. The subsequent steps are carried out by phosphoribosylformimino-5-aminoimidazole carboxamide ribotide isomerase [EC: 5.3.1.16], imidazole glycerol-phosphate synthase [EC:4.3.2.10], imidazoleglycerol-phosphate dehydratase [EC: 4.2.1.19], histidinol-phosphate aminotransferase [EC: 2.6.1.9], histidinol-phosphate phosphatase [EC: 3.1.3.15] and histidinol dehydrogenase [EC: 1.1.1.23] [6, 69].

The histidine pathway forms an essential connection between amino acid, purine, and thiamine biosynthesis [71]. It is one of the most highly regulated amino acid biosynthesis systems due to also being one of the most energetically consuming anabolic pathways [72]. The cost of synthesising one molecule of histidine (when histidine is not present in the growth media) was determined to be around 41 molecules of ATP in *Salmonella typhimurium* [73] and therefore stringent regulation of the pathway is essential to conserve cell resources. The initial step in the pathway (coded for by the *hisG* gene) is highly regulated by allosteric feedback inhibition by histidine. Some organisms also have initial step competitive orthosteric inhibition by substrate analogues such as AMP and ADP [74], product inhibition from PRATP [75] and positive cooperative inhibition by guanosine tetraphosphate (ppGpp) when also in the presence of histidine [76]. Interestingly, guanosine tetraphosphate accumulates when bacteria are starved for amino acids and is a positive effector of *his* operon transcription [71]. Therefore, ppGpp is an inhibitor of ATPPRT (the first committed step of the pathway) in the presence of histidine, but a transcriptional effector in the absence of histidine.

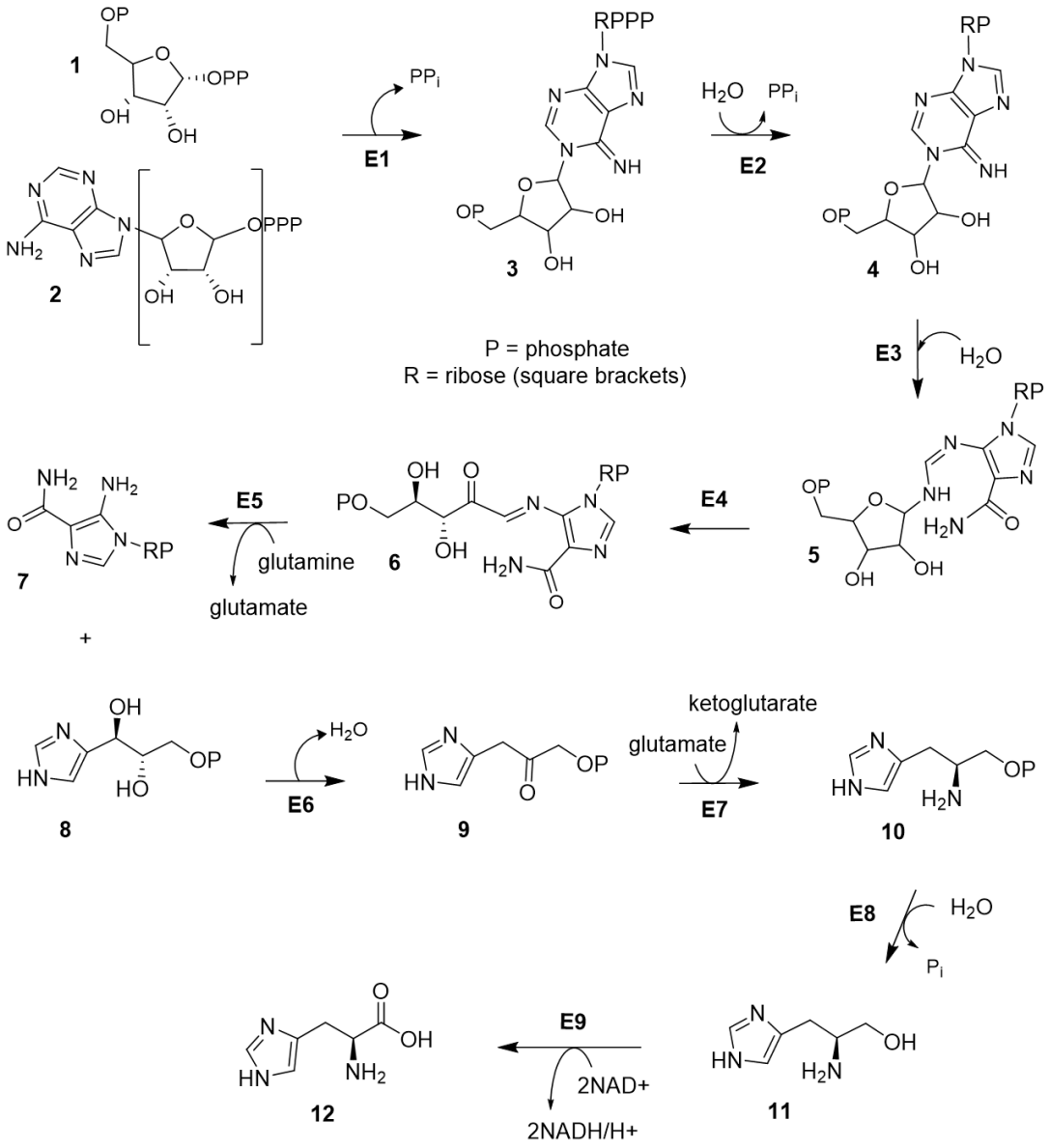


Figure 1.7: Histidine Biosynthetic pathway [6]. Enzymes: ATP phosphoribosyltransferase (E1), phosphoribosyl ATP pyrophosphohydrolase (E2) phosphoribosyl AMP cyclohydrolase (E3) phosphoribosylformimino-5-aminoimidazole carboxamide ribotide isomerase (E4), imidazole glycerol-phosphate synthase (E5), imidazoleglycerol-phosphate dehydratase (E6), histidinol-phosphate aminotransferase (E7), histidinol-phosphatase (E8) and histidinol dehydrogenase (E9). Compounds: PRPP (1), ATP (2), PRATP (3), phosphoribosyl-AMP (4), phosphoribosyl-formimino-AICAR-phosphate (5), phosphoribulosyl-formimino-AICAR-phosphate (6), 5-aminoimidazole-4-carboxamide ribonucleotide (AICAR) (7), imidazole-glycerol-phosphate (8), imidazole-acetol-phosphate (9), L-histidinol-phosphate (10), L-histidinol (11), L-histidine (12).

In general, bacteria grown in minimal-glucose medium has four-fold greater *his* operon expression than bacteria grown in rich medium and thus the growth rate of the cell is related to the metabolic regulation of the cell in response to the overall amino-acid

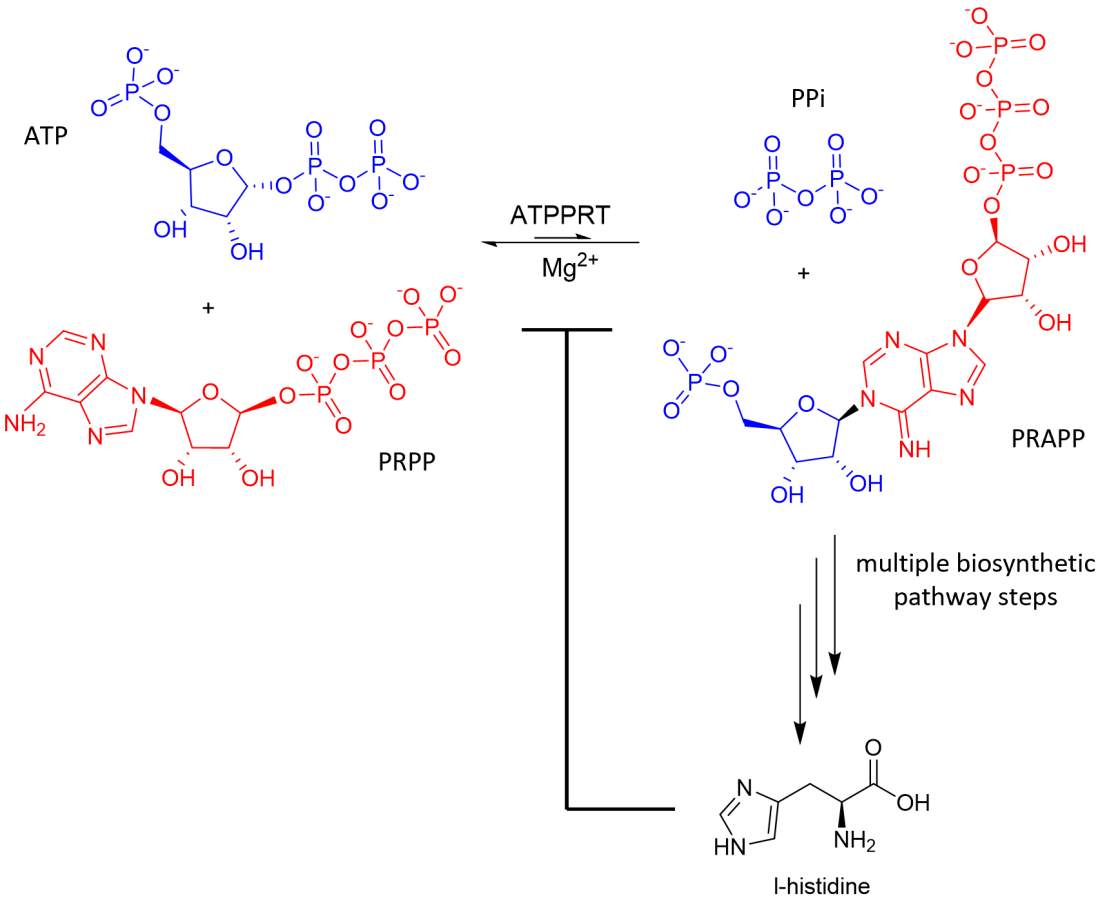


Figure 1.8: ATPPRT-catalysed reaction is inhibited by the product of the pathway, histidine [70].

supply in the cell [77]. In minimal medium conditions, the rate of histidine biosynthesis is chiefly controlled by regulation of HisG enzymatic activity [71]. Since histidine biosynthesis is regulated mainly by modulating the flow of intermediates through the pathway [71], this array of inhibiting molecules of the initial step allows delicate control of the rate of histidine biosynthesis across a variety of cellular metabolic states. However, histidine biosynthesis is also regulated through *his* operon expression. The *his* operon is comprised of a cluster of structural genes required for histidine biosynthesis (Figure 1.9[78]) and its expression is upregulated when the bacteria is starved of histidine [79]. It is generally accepted that two mechanisms at the transcription level are responsible for the regulation of the *his* operon [71]. The first regards transcription ignition by the textithis operon primary promoter (*hisp1*) through concentration change of ppGpp. The second regards transcription of the *his* structural genes by an attenuation mechanism vigilant of intracellular concentrations of His-tRNA^{His}. His-tRNA^{His} concentrations fluctuate in response to cellular histidine concentration, histidyl-tRNA synthetase and chromosomal DNA supercoiling levels in response to anaerobiosis and osmolarity [71].

Histidine insensitive mutants of $HisG_L$ have previously been prepared to increase histidine production for industrial and biochemical uses in strains such as *Corynebacterium glutamicum*, *Brevibacterium flavum*, *Serratia marcescens* and *Escherichia coli*[80]. Initially, basic cloning techniques to isolate mutants with histidine insensitivity were utilised for industrial strain preparation, for example A PRT deficient mutant of *C. glutamicum* was shown to produce twice as much histidine by batch fermentation as the wild type[81]. More recently, specific residue mutations (N215K/L231F/T235A) designed to disrupt histidine binding to ATPPRT in *C. glutamicum* resulted in 37-fold less histidine sensitivity than the wild type with no appreciable change to enzyme activity [82]. These or similar strains are still in use [83] but histidine sensitivity is not the only modification available to improve histidine production in bacterial strains. An example of an alternative approach to improved histidine yield utilised His_{E271K}, an *E.coli* strain insensitive to histidine but susceptible to inhibition by AMP, ADP and AICAR. By *purH* and *purA* gene overexpression and a modified PitA-dependent phosphate/metal transport system (via a pitA gene deletion), a new *E. coli* strain capable of increased histidine accumulation was developed[80].

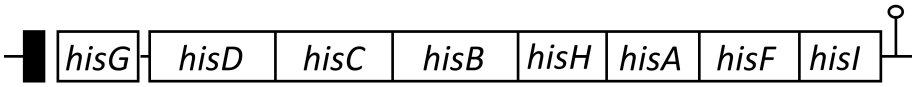


Figure 1.9: Schematic representation of the organisation of the histidine biosynthetic genes in *E. coli* and *Salmonella typhimurium* bacteria. The schematic is labelled as follows. Black box depicts leader sequence and white boxes depict the structural gene. Lines between genes indicate intergenic regions [78].

1.2.2 ATP phosphoribosyltransferases

ATPPRT is part of the phosphoribosyltransferase (PRT) superfamily which mediate reactions involving the transfer of phosphoribosyl groups to nitrogen moieties[72]. ATPPRT specifically transfers the phosphoribosyl moiety of PRPP to ATP to produce PRATP [70].

Long- and Short-Form ATPPRTs

ATPPRTs appear in one of two genetically distinct forms. $HisG_L$ is termed the 'long form', is homo-hexameric and around 300 amino acids in length whereas $HisG_S$ is shorter at around 200 amino acids [84] and is a homodimer with each subunit comprising of a C-terminus regulatory domain and a catalytic domain. The discrepancy in chain length is around 100 residues found at the C terminus in the long form but not in the short form [85], which is responsible for allosteric inhibition of $HisG_L$ [86]. $HisG_S$ and $HisG_L$ share

~25% sequence identity [70] and these regions of similarity are found in the homologous catalytic domains [87].

The long-form is a fully catalytically active enzyme [69] encoded by the *hisG* gene and is the most studied form of the enzyme [86]. The 'short-form'is also encoded by the *hisG* gene but a different variant to $HisG_L$, one which lacks the C-terminal domain. $HisG_S$ is catalytically active but insensitive to histidine [88, 89] and requires allosteric activation by its regulatory protein, HisZ (encoded by the *hisZ* gene) to reach maximum catalytic activity and regulate histidine feedback inhibition [85]. The short form ATPPRT holoenzyme exists as a hetero-octameric protein comprising of two $HisG_S$ dimers and a HisZ tetramer [20, 84, 90]. The four HisZ subunits sit as a cross-shaped tetramer in the middle of the structure with two $HisG_S$ dimers flanking top and bottom (Figure 1.10). HisZ is solely a regulatory protein. It does not contain a catalytic domain (so is inactive on its own), is not found in species containing $HisG_L$ or in species which cannot synthesise histidine [84].

Until recently it was thought that $HisG_S$ was only able to catalyse the ATPPRT reaction after forming the non-covalent octameric complex with regulatory protein HisZ [87]. However, at least in the cases of *Psychrobacter arcticus* [20] and *Lactococcus lactis* (*L.lactis*) [91], $HisG_S$ can catalyse the reaction alone but at considerably lower rates than in conjunction with HisZ. Alone, $HisG_S$ forms homodimers that are insensitive to histidine but when in complex with HisZ, the ATPPRT is histidine sensitive [20, 91]. HisZ is responsible for the histidine feedback inhibition and is therefore responsible for both activation and inhibition of short-form ATPPRTs.

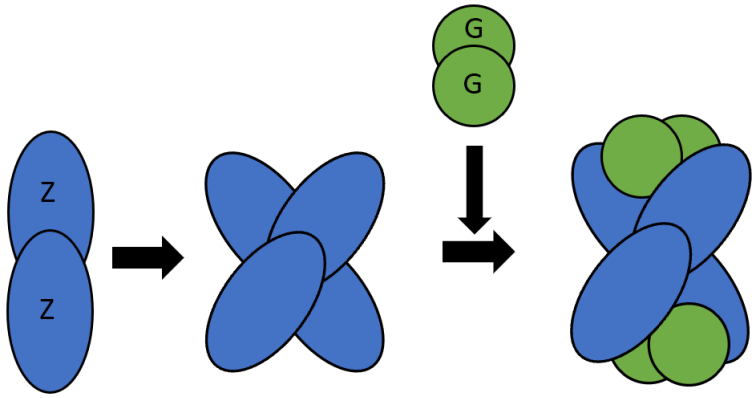


Figure 1.10: Schematic representation of short form ATPPRT structure formation from two HisZ dimers and two $HisG_S$ dimers.

Evolution of ATPPRT

Whether an organism has a short or long form of HisG is based on the organism's evolution. The short form is most commonly found in archaea but also in some bacteria (including *S. aureus* and *P. arcticus*) and the long form is most commonly found in bacteria (such as *E. coli*) and histidine-producing eukaryotes [58, 68, 86, 92–94]. Both $HisG_L$ and $HisG_S$/HisZ (the ATPPRT apoenzyme) have greatly reduced activity in the presence of histidine due to feedback inhibition [69].

$HisG_S$ evolved separately [84] from HisZ, but also separately from other classes of PRT. The only protein with clear evidence of a shared ancestor to $HisG_S$ is $HisG_L$ [84]. They are both likely to have descended from a common ancestor with the periplasmic binding protein (PBP II) family which includes phosphate and sulphate binding proteins (although none of the PBP IIs show PRT function) [95].

HisZ and histidyl-tRNA synthetase (HisRS) are paralogs [70, 85] and so are encoded by genes that are derived from the same ancestral gene. HisZ does not have aminoacylation activity (or any other catalytic activity) but like HisRS, it contains the similar histidine-binding motifs characteristic of tRNA synthetases [87]. It is most likely that ATP and tRNA binding were removed during the evolutionary path. Some HisZs have greater similarity to HisRS than others, for example, there are a subset of short form ATPPRTs whose HisZ subunits lack a C-terminal domain found in many other HisZs and HisRS. HisZs with this C-terminal domain have similar regulatory properties to those without and are still histidine sensitive [86] so it is unclear what specialised function, if any, these extra residues are responsible for.

Allostery of ATPPRTs

Previous allosteric study of ATPPRTs has focused on the long form of the enzyme. There is currently no clear consensus on the exact mechanism of allostery in ATPPRT. There are two proposed mechanisms, one where ligand binding causes a large conformational change stabilizing one oligomeric state and another suggesting no change in oligomeric order but local conformational changes capable of tightening the overall structure. In the case of *Mycobacterium tuberculosis*, interconversion of dimer-hexamer states does not account for the allosteric inhibition seen in the presence of histidine [69] and is more likely a shift between two hexametric states varying by a domain III shift, an overall tightening and subtle conformational changes upon binding of histidine[94]. Allosteric histidine binding is thought to trap the enzyme in a form that is inactive but still capable of catalysis, since upon increasing concentrations of histidine, a decrease in the amplitude of the burst phase in pre-steady state kinetics is observed [58]. The presence of a burst phase indicates that product release is rate limiting for *M. tuberculosis* $HisG_L$.

In the case of *C. jejuni* however, there is evidence from computational and mutagenesis approaches to suggest that a hinge twist within the flexible catalytic domain creates distinct conformational ensembles. These population shifts have varying dynamic properties and are capable of facilitating the allosteric response[96]. On the basis of these two key studies, there is suggestion that the latter proposed mechanism of small localised changes is more likely to be the underlying allostery mechanism in ATPPRTs than a large conformation change to stabilise one oligomeric state over another. Further analysis of ATPPRTs from a wider array of organisms is required to determine the validity of this statement.

1.2.3 *Psychrobacter arcticus*

P. arcticus, a gram-negative, rod-shaped, heterotrophic, psychrophilic bacterium from the Siberian permafrost [97] which grows between at -10 and 28°C. *P. arcticus* was the first cold-adapted bacterium to have its genome sequenced and has been used as a cold-adapted model [97–99] due to its abundance in permafrost and to its growth condition in the wild being always under sub-zero temperatures. Compared to mesophilic bacteria, psychrophiles can have differences in membrane composition, contain cold shock proteins[97], better utilise acetate as a carbon source[99], have a lower frequency of amino acids such as proline and arginine in their proteins [100] and have increased protein flexibility at low temperature, at least in the outer shells of the protein. *P. arcticus* contains a short-form ATPPRT, as does *Staphylococcus aureus* and *Acinetobacter baumannii*, both of which urgently require new antibiotics. Study of *P. arcticus* ATPPRT (which is regulated by both allosteric inhibition and activation) is therefore useful as a model to provide insight into a valuable drug target but also allosteric intricacies. *P. arcticus* has been studied extensively in the da Silva lab [20, 101, 102] to shed light on short-form ATPPRT kinetics, structure, mechanism and allostery with a view to improve histidine biosynthesis industrially. All purification protocols are already established and so in the interest of time and resources, *Pa*ATPPRT was used for the present study.

1.2.4 Structure of *Psychrobacter arcticus* ATPPRT

Similarly to *Thermotoga maritima* ATPPRT [84] and *Lactococcus lactis* ATPPRT [103], the *Pa*ATPPRT apoenzyme is a hetero-octamer comprising of two $HisG_S$ dimers flanking top and bottom of a HisZ tetramer which sits crossed, in a head-to-tail arrangement, at the centre (Figure 1.11) [20]. The $HisG_S$ subunits have two domains [20] which is in line with *Thermotoga maritima* ATPPRT [84] and *Lactococcus lactis* ATPPRT [70] but also $HisG_L$ catalytic domains [104, 105]. The HisZ subunits contain a large N-terminal domain and smaller C-terminal domain [20]. This is the same as HisRS but not *Thermotoga maritima*

Figure 1.11: *Pa*ATPPRT quaternary structure. Four HisZ subunits (yellow, blue, cyan and green) sit crossed in the centre of the hetero-octamer with two pairs of $HisG_S$ subunits (red and orange) flanking top and bottom.

ATPPRT [84] or *Lactococcus lactis* ATPPRT [70]. The latter two have not retained the C-terminal domain during evolution.

Structure of *Psychrobacter arcticus* ATPPRT:PRPP

The active site of *Pa*ATPPRT is very similar to that of *Pa*$HisG_S$ alone and the other characterised ATPPRTs [72, 86, 106–108]. The structures of both *Pa*$HisG_S$ and *Pa*ATPPRT bound to PRPP, PRPP:ATP, and PRATP have been published [101]. No *Pa*ATPPRT structure with ATP (or its alternative substrate ADP) bound in the active site could be obtained which suggests PRPP may be required to bind first.

When PRPP is bound to *Pa*ATPPRT the substrate is secured to the active site via main and side chain interactions between the PRPP binding loop and the phosphate tail of PRPP. There is also a hydrogen bond (H-bond) that stops PRPP from blocking the ATP binding site between the substrate 2-OH group and side chain residue Asp176. This is similar to the PRPP bound structure of *L. lactis* [107] and *Pa*$HisG_S$ alone, but not *C. jejuni* [86].

Structure of *Psychrobacter arcticus* ATPPRT:PRPP:ATP

The *Pa*ATPPRT:PRPP:ATP bound structure is similar to *Pa*ATPPRT:PRPP but the Asp32-Asp176 H-bond has been broken by the Arg 32 side chain rotating 180° around it's cγ-cδ axis with its guanidinium group now sitting parallel to the adenine ring. An Arg132-Arg179 salt bridge keeps the ATP site clear (Figure 1.12, B). This is found in the *Cj*ATPPRT :ATP bound structure[106] and is a conserved residue across ATPPRTs. Asp94 H-bonds to the 2'OH and 3'OH of ATP (also true of *C. jejuni* and *Mycobacterium tuberculosis (M. tuberculosis)* [86, 106] and a Lys137 salt bridge with beta-PO_4.

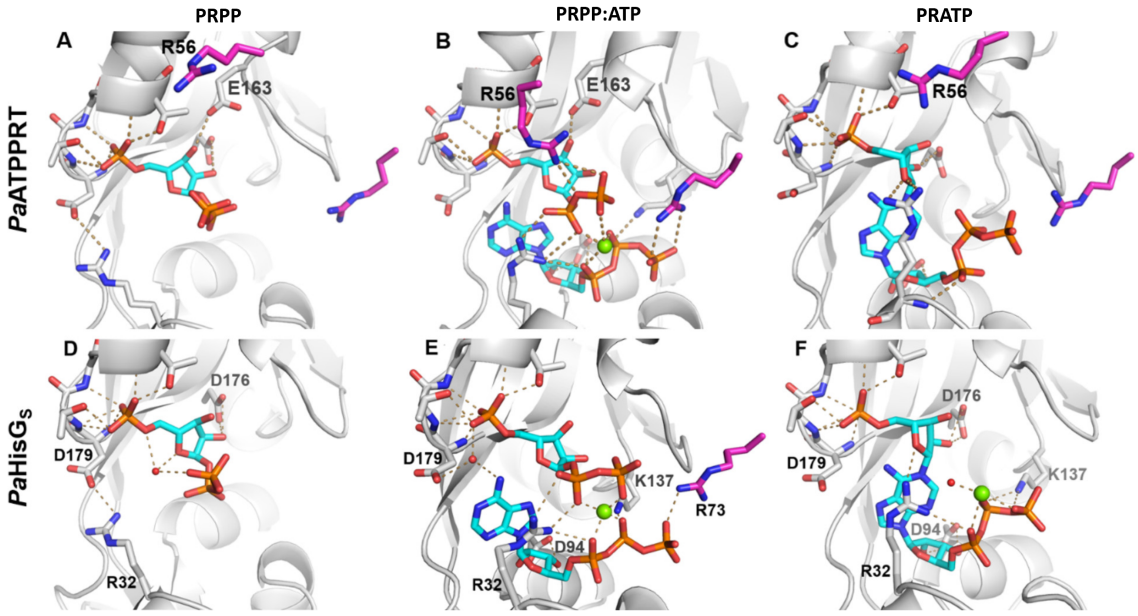


Figure 1.12: Enlarged view of the active-site of *Pa*ATPPRT bound to (A) PRPP, (B) PRPP:ATP, (C) PRATP and of *Pa*HisG$_S$ bound to (D) PRPP, (E) PRPP:ATP and (F) PRATP. The figure shows polar interactions (grey dashed lines), magnesium ions (green spheres), nitrogen atoms (blue), oxygen atoms (red), phosphorus (orange), ligands (stick models carbon atoms are cyan), protein backbone (grey ribbon diagram with selected residues from main and side chains shown as stick models, carbon atoms are grey), side chains of two Arg residues contributed by the adjacent subunit (stick models with carbon atoms in magenta). Residues whose interactions with ligands are absent in the *Pa*HisG$_S$ structures are labelled in the *Pa*ATPPRT structures and vice versa [101].

There is further stabilisation from a salt bridge between the γ-PO_4 and Arg73 on an adjacent *Pa*HisG$_S$ subunit, but this may not be significant as a comparable *L. lactis* mutant, lacking the capability to form the salt bridge, had no significant binding differences. ATP and PRPP are brought into proximity for reaction by a salt bridge between Arg32 and α-PO_4 groups of both substrates. Additionally, the electrostatic repulsion from the negatively charged phosphate groups is reduced by a Mg_{2+} ion coordinated between the α and β-PO_4 of both substrates. Arg32 is the only residue to

interact with the PP_i moiety – it is likely that another Mg^{2+} ion is required to stabilize the PP_i leaving group but this was not observed in the crystal structure. This ternary complex structure has not converted to product because the adenine ring is bound with the N1 ATP facing away from the PRPP C1 (the site of nucleophilic attack). This anticatalytic rotamer could be reason this structure was trapped with both products in the active site but it could also be because the reaction favours the reactants. In solution the necessary orientation for catalysis would be able to be sampled but the crystal form probably lacks the flexibility. This is not unique to *Pa*ATPPRT however, *C. jejuni* also showed electron density for both substrates in the active site [86].

In a ternary complex with its alternative substrate ADP, ADP mimics the orientation of ATP in the binding site. Almost all interactions except the Arg73 salt bridge are preserved, including coordination to Mg^{2+}. The Arg73 on the neighbouring subunit is in the same position but is too far away to reach beta-PO_4 of ADP. Considering PRPP:ADP:*Pa*$HisG_S$ and PRPP:ADP:*Pa*ATPPRT (Figure 1.13) structures were obtained where the ADP binds in a similar mode as ATP does. This suggests that ADP may be a substrate replacing ATP, and not an inhibitor as is the case with $HisG_L$-based ATPPRTs [109] since it can bind in the same active site with similar contacts.

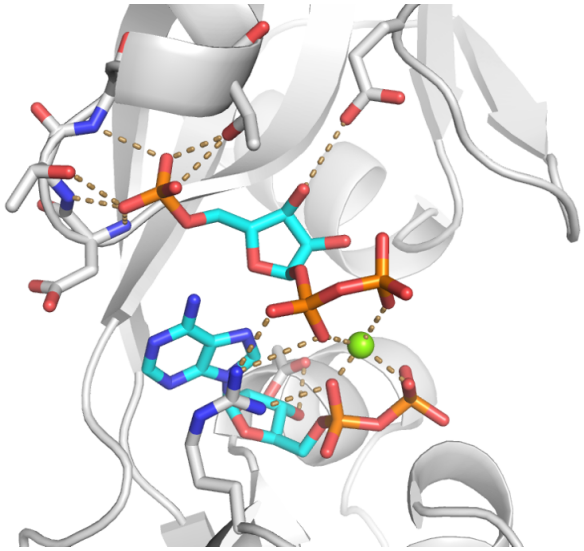


Figure 1.13: Enlarged view of active-site of *Pa*ATPPRT bound to PRPP:ADP. The figure also shows magnesium (green spheres) nitrogen atoms (blue), oxygen atoms (red), phosphorus atoms (orange), ligands (stick models, carbon in cyan), protein backbone (ribbon diagram with selected main and side chain residues as stick models, carbon in grey) [101].

In the *Pa*ATPPRT :PRPP bound structure and the *Pa*ATPPRT :PRPP:ATP bound there is an additional H-bond between Glu163 and the 3-OH group (which likely helps position the substrate for catalysis) that is not seen in the solitary *Pa*$HisG_S$:PRPP bound structure. The PP_i moiety of the PRPP is very flexible and has no observed specific interaction

with the protein, only one interaction overall – with a water molecule that also interacts with the less flexible 5-phosphoribosyl moiety [101]. Structure of *Psychrobacter arcticus* ATPPRT:PRATP

In the *Pa*ATPPRT :PRATP bound structure, density for the entire product was observed (there is a similar structure for *C. jejuni* which also shows product electron density[86]). However, the Mg^{2+} ion density could not be definitively assigned but it is thought to be similarly to *Pa*HisG$_S$ alone whereas in the PRATP bound structure the Mg^{2+} ion is coordinated to the α- β- and γ-PO_4 of PRATP. Overlay showed good overlap between the Mg^{2+}, phosphoribosyl and triphosphoribosyl moiety position in all structures. There is however a 1.6 Å shift of the η-NH group of Arg32 towards the PRPP binding site. There is still a parallel relationship between the adenine ring and the guanidinium, but a H-bond is now donated to O4" of PRATP.

1.2.5 Activation and Mechanism of *Pa*ATPPRT

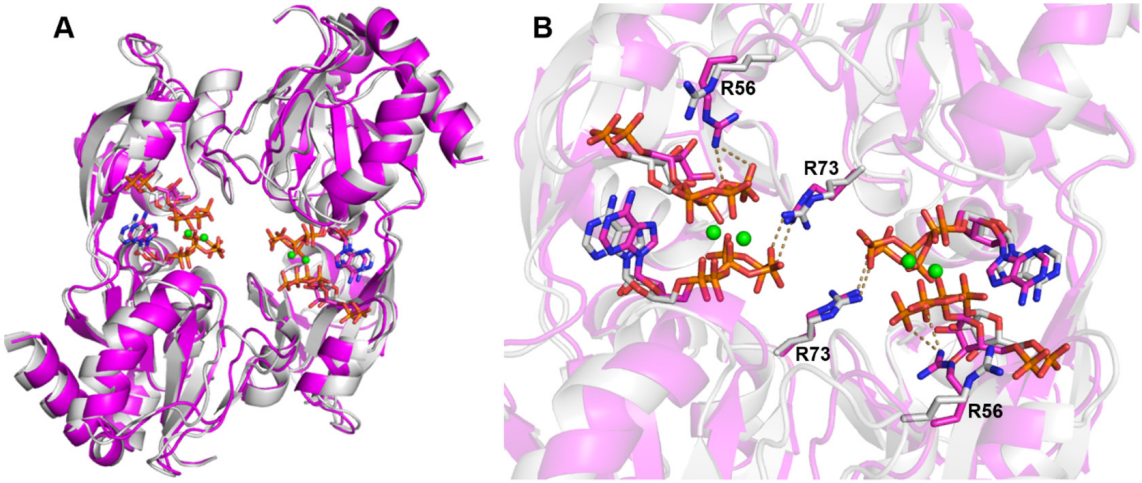


Figure 1.14: Overlay of *Pa*HisG$_S$ dimers from *Pa*ATPPRT and *Pa*HisG$_S$ alone bound to PRPP, ATP and Mg^{2+}. (A) Overlay of *Pa*HisG$_S$ dimers of non-activated *Pa*HisG$_S$ (grey) and activated PaATPPRT (magenta). (B) Enlarged view of overlay comparison of non-activated and activated *Pa*ATPPRT active sites showing position change of Arg56 and Arg73 residues in PRPP:ATP bound structures. Polar interactions are shown by dashed lines, magnesium ions (green spheres), non-activated *Pa*HisG$_S$ (grey) and activated *Pa*ATPPRT (magenta), Arg56 and Arg73 side chains (stick models with nitrogen in blue, oxygen in red, phosphorus in orange, and carbon in magenta (*Pa*ATPPRT) and grey (*Pa*HisG$_S$).

There is little variation in structure between *Pa*HisG$_S$ non-activated) and *Pa*ATPPRT (activated) despite the presence of HisZ leading to a large activation of *Pa*HisG$_S$ [101] (Figure 1.14, A). There is an observed tightening of the overall *Pa*HisG$_S$ dimer upon HisZ binding which increases the cross-dimer contacts of some side chains (Figure 1.14,B). The

structures of both *Pa*HisG$_S$ (non-activated) and *Pa*ATPPRT (activated) bound to PRPP, PRPP:ATP, and PRATP are similar, with substantial differences only manifested in the PRPP:ATP bound structures (Figure 1.12).

Based on these structures [101] it is proposed that PRPP binds to the enzyme first before ATP binding can occur. This is incongruent with published long form ATPPRT HisG$_L$ structures *C. jejuni* and *M. tuberculosis* [86, 106] (which shows ATP binding first in an ordered Bi-Bi kinetic mechanism[110] but supported by the only other short form ATPPRT:PRPP complex published: Ł.lactis[91]. Product release, however, follows the same pattern as both long and short form ATPPRTs previously published [86, 91, 107] with PP$_i$ leaving first, followed by PRATP. This hypothesis has been corroborated by kinetic interrogation [102] Kinetic analysis by initial velocity studies with both substrates determined *Pa*HisG$_S$ alone forms a ternary complex with both substrates in the active site congruently and that binding is sequential. Binding of PRPP to *Pa*HisG$_S$ was detected by isothermal calorimetry (ITC) whereas binding of ATP to PaATPPRT could be not be established. Similarly, differential scanning fluorometry (DSF) detected stabilisation of T$_m$ upon PRPP binding but not ATP. Given the extensive similarity between *Pa*HisG crystal structures and *Pa*ATPPRT, the mechanism is assumed to be the same for both. The kinetic mechanism of *Pa*HisG$_S$ and *Pa*ATPPRT is summarised in Figure 1.15.

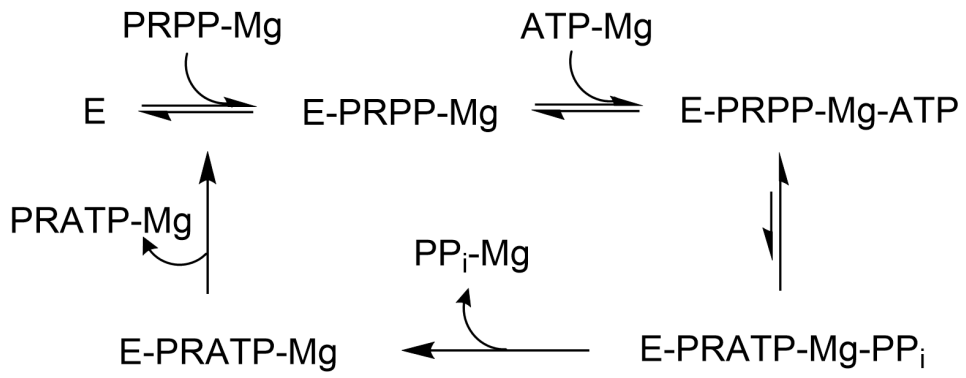


Figure 1.15: Proposed reaction scheme of *Pa*ATPPRT.

Replacing Mg^{2+} in the buffer with Mg^{2+} does not affect the rate of *Pa*ATPPRT reaction therefore the rate limiting step is unlikely to be the stabilisation of phosphate groups. This is the case however for *Pa*HisG$_S$, which showed an increase in rate when Mn^{2+} was used. A stronger Lewis acid stabilises charges more efficiently, leading to a faster rate of PRATP production. The rate limiting step for *Pa*HisG$_S$ is therefore the interconversion between the *Pa*HisG$_S$:PRPP:ATP and *Pa*HisG$_S$:PRATP:PP$_i$ ternary complexes. The rate limiting step for *Pa*ATPPRT catalysis is product release (as determined by burst kinetics)[102].

1.3 Potential Impact

The current work investigates the mechanism of histidine inhibition in *Pa*ATPPRT through steady-state and pre-steady-state kinetics, thermal-shift assays, and X-ray crystallography. Inhibition by AMP and activity with ADP are also characterised. These findings provide the best characterisation of inhibition in a $HisG_S$-based ATPPRT, and may be useful as a model to understand disease-relevant ATPPRTs. For example, *Acinetobacter baumannii*, which belongs to the same bacterial family as *P. acticus* (Moraxellaceae), is at the top of the World Health Organisation (WHO) list of the "dirty dozen", a list of 12 bacterial species against which novel antimicrobials are desperately needed [111]. *A. baumannii* has a short-form ATPPRT of which $HisG_S$ and HisZ share 70% and 43% sequence identity with their *P. arcticus* counterparts, respectively. Recently, HisZ has been shown to be essential for *A. baumanii* survival, and $HisG_S$ essential for *A. baumannii* persistence in the lung, which is a requirement for virulence in this species[111]. Therefore, targeting ATPPRT could be a strategy to thwart *A. baumanii* infection, and given the high similarity between *A. baumanii* and *P. arcticus* ATPPRTs, the mechanistic inferences from this work could aid drug design. Furthermore, *in vivo* histidine production without feedback inhibition would be a lucrative avenue for amino acid production. Therefore, the current work not only delivers insight into inhibition of short-form ATPPRTs but provides relevant information for biotechnology approaches to histidine synthesis and antimicrobial discovery.

Chapter 2

MATERIALS AND METHODS

2.1 Materials

2.1.1 Reagents

ATP, PRPP, adenosine 5'-triphosphate (ADP), L-histidine, 3-(2-thienyl)-L-alanine (TIH), $MgCl_2$, dithiothreitol (DTT), and tricine were purchased from Sigma-Aldrich. HEPES and NaCl were sourced from Formedium. EDTA-free Cϕmplete Protease Inhibitor Cocktail was purchased from Roche and Bug Buster, from Merck. A Gibson Assembly kit was purchased from New England Biolabs. All primers were from IDT and all commercially available enzymes were from Thermo Fisher Scientific or New England Biolabs. *Mycobacterium tuberculosis* inorganic pyrophosphatase *(Mt*PPase) was purified in-house as previously published [20]. Unless otherwise stated, all other reagents were purchased from readily available commercial sources and all were used without further purification. Expression vectors and g-block DNA were ordered from DNA 2.0.

2.1.2 Materials and Equipment

Ultrafiltration membranes were manufactured by Millipore and dialysis cassettes were Slide-A-Lyser^{TM} purchased from Thermo Fisher Scientific. Centrifugal filter units were manufactured by Amicon. The high-pressure cell disruptor used for cell lysis was produced by Constant Systems. All protein purifications were carried out using an AKTA Start FPLC system and all nickel columns were purchased from GE Lifesciences. Kinetic experiments were carried out on either a UV-2600 spectrophotometer outfitted with a CPS unit for temperature control (Shimadzu) or an SX-20 stopped-flow spectrophotometer fitted with a 5 μL mixing cell (0.5-cm path length and 0.9-ms dead-time) by Applied Photophysics. Unless otherwise stated, all other materials were

purchased from readily available commercial suppliers.

2.1.3 Services

All mass spectrometry analysis was carried out by the BSRC Mass Spectrometry Facility at the University of St Andrews. All DNA sequencing was undertaken by Eurofins Genomics using their Mix2Seq kit.

2.1.4 Plasmid, Protein and Primer Sequences

Plasmid Sequences

Expression vectors pJexpress414 and pJexpress411 containing the DNA sequences encoding for *Sa*HisZ and *Sa*HisG$_S$, respectively, with an N terminus 6-histidine tag (6xHis-tag) and tobacco etch virus protease (TEVP) cleavage site were ordered from DNA 2.0 with codon-optimisation for expression in *Escherichia coli*.

Expression vectors pJexpress414 and pJexpress431 containing the DNA sequences encoding for *Pa*HisZ and *Pa*HisG$_S$, respectively, with an N terminus 6-histidine tag (6xHis-tag) and TEVP cleavage site were ordered from DNA 2.0 with codon-optimisation for expression in *E. coli*.

Primer Sequences

*Sa*HisG$_S$ MBP Primer 1 Reverse
5′- TCG TAT TAA TTT CGC GGA AGG -3′

*Sa*HisG$_S$ MBP Primer 2 Forward
5′- CT GTT GAG CGT TGA TAA C -3′

T7 Promoter Primer
5′-TAATACGACTCACTATAGGG-3′

Internal Sequencing Primer
5′-CAT TTA AAG GCC AGC CTA GCA-3′

Maltose Binding Protein (MBP) G-block

CCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCC
CTCTAGAAATAATTTTGTTTAACTTTTAAGGAGGTAAACATATGCACCACCATCA
TCATCACATGAAAATTAAAACAGGTGCCCGCATTCTGGCCTTGTCTGCGTTGACG
ACCATGATGTTTAGCGCATCAGCATTAGCGAAAATTGAAGAAGGTAAGCTGGTC
ATCTGGATCAACGGCGATAAAGGCTACAACGGCTTGGCAGAAGTTGGCAAAAAA
TTTGAAAAAGATACCGGTATTAAAGTCACCGTGGAGCACCCTGATAAACTTGAA
GAGAAGTTTCCACAAGTCGCCGCGACCGGCGACGGGCCCGATATTATTTTTTGG
GCACATGATCGTTTTGGCGGATATGCACAATCGGGACTTCTGGCCGAAATTACA
CCAGACAAGGCATTCCAGGATAAATTATACCCGTTTACATGGGATGCTGTACGG
TATAATGGCAAACTTATCGCGTATCCAATTGCGGTTGAGGCGCTGTCGTTAATTT
ATAACAAAGATCTCCTGCCGAACCCGCCTAAAACCTGGGAGGAGATTCCTGCAC
TGGACAAGGAGCTCAAAGCAAAAGGCAAGTCGGCACTGATGTTTAATCTTCAGG
AACCGTACTTTACGTGGCCGCTGATTGCTGCAGACGGCGGTTATGCGTTTAAATA
TGAGAATGGAAAGTATGATATTAAAGATGTTGGTGTAGATAACGCTGGCGCTAA
AGCGGGTCTGACATTCTTGGTAGATCTGATTAAAAACAAACACATGAACGCGGA
TACGGATTACTCTATCGCCGAAGCCGCCTTCAACAAGGGCGAAACAGCAATGAC
CATCAATGGGCCGTGGGCATGGTCAAACATTGACACGAGCAAGGTGAATTATGG
CGTGACGGTGCTGCCGACATTTAAAGGCCAGCCTAGCAAACCCTTTGTTGGCGTT
CTGTCAGCAGGTATCAATGCGGCTAGCCCTAATAAAGAACTGGCGAAAGAATTT
CTGGAGAACTATCTTTTAACGGATGAAGGGCTGGAAGCCGTGAACAAAGATAAA
CCCCTGGGTGCCGTCGCCCTGAAATCATACGAAGAGGAATTAGTGAAAGACCCG
CGCATCGCAGCGACCATGGAAAACGCCCAAAAAGGTGAGATCATGCCGAACATT
CCCCAGATGTCTGCCTTTTGGTATGCCGTACGCACCGCGGTTATCAATGCTGCAT
CAGGTCGCCAGACCGTTGATGAAGCCCTGAAGGATGCCCAGACCGAGAATCTGT
ACTTTCAAGGTATGCTGCGTATCGCTATTGCGAAGGGCCGTCTGATGGACAGCCT
GATCAATTATCTGGACGTCATTGAGTACACCACGTTGTCCGAAACGTTAAAGAAT
CGCGAACGCCAGCTGCTGTTGAGCGTTGATA

Protein Sequences

Mass and extinction coefficients at 280 nm (ε280) expressed proteins were calculated by the software ProtParam. The 6xHis tagged proteins were cleaved by TEV protease using the recognition sequence ENLYFQ\G, where \ denotes the cleavage site. The sequences presented below also contain a \ to denote the TEV Protease cleavage site.

6H*Sa*HisZ (33460 Da, ε280 = 47330 M^{-1} cm^{-1})

MHHHHHHENLYFQ\GMNNSEQLIALKESETAFLKYFNKADYELVDFSVVEKLDWKQ
LNHEDLQQMGERNFWQHEHQIYALRNDFTDQLLRYYSMYPTAATKVAYTGLIIRNN
EAAVQVGLENYAPSLANVQQSLKLFIQFIQQQLRDNVHFVVLGHYQLLDALLDKSLQ
TPDILSMIEERNLSGLVTYLSTEHPIVQILKENTQQQLNVLEHYIPNDHPALVELKIWE
RWLHTQGYKDIHLDITAQPPRSYYTGLFIQCHFAENESRVLTGGYYKGSIEGFGLGLTL

*Sa*HisZ (31711 Da, ε280 = 45840 M^{-1} cm^{-1})

GMNNSEQLIALKESETAFLKYFNKADYELVDFSVVEKLDWKQLNHEDLQQMGERNF
WQHEHQIYALRNDFTDQLLRYYSMYPTAATKVAYTGLIIRNNEAAVQVGLENYAPSL
ANVQQSLKLFIQFIQQQLRDNVHFVVLGHYQLLDALLDKSLQTPDILSMIEERNLSGL
VTYLSTEHPIVQILKENTQQQLNVLEHYIPNDHPALVELKIWERWLHTQGYKDIHLDI
TAQPPRSYYTGLFIQCHFAENESRVLTGGYYKGSIEGFGLGLTL

6H*Sa*HisG$_S$ (24403 Da, ε280 = 13410 M^{-1} cm^{-1})

MHHHHHHENLYFQ\GMLRIAIAKGRLMDSLINYLDVIEYTTLSETLKNRERQLLLSVD
NIECILVKGSDVPIYVEQGMADIGIVGSDILDERQYNVNNLLNMPFGACHFAVAAKPE
TTNYRKIATSYVHTAETYFKSKGIDVELIKLNGSVELAGVVDMVDGIVDIVQTGTTLK
ANGLVEKQHISDINARLITNKAAYFKKSQLIEQFIRSLEVSIANA

6H*Sa*HisG$_S$-MBP (67302 Da, ε280 = 79760 M^{-1} cm^{-1})

MHHHHHHMKIKTGARILALSALTTMMFSASALAKIEEGKLVIWINGDKGYNGLAEV
GKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITP
DKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKEL
KAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFL
VDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFK
GQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEE
ELVKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQ
TENLYFQ\GMLRIAIAKGRLMDSLINYLDVIEYTTLSETLKNRERQLLLSVDNIECILVK
GSDVPIYVEQGMADIGIVGSDILDERQYNVNNLLNMPFGACHFAVAAKPETTNYRKI
ATSYVHTAETYFKSKGIDVELIKLNGSVELAGVVDMVDGIVDIVQTGTTLKANGLVEK
QHISDINARLITNKAAYFKKSQLIEQFIRSLEVSIANA

6H*Pa*HisZ (44703 Da, ε280 = 27390 M^{-1} cm^{-1})

MHHHHHHENLYFQ\GMLPDGVADVLFEDAHKQEVLRHQLTQQLITHGYQLVSPPM

IEFTESLLSGASEDLKRQTFKIIDQLTGRLMGIRADITPQILRIDAHHGGDGIARYCYAG
DVIHTLPSGLFGSRTPLQLGAEIFGCESIAADIELIDVLFSMINSLDMSAVLHVDLGHVT
IFKRLAELAALSASDTEQLMQLYANKNLPELKQVCQVLPMGSDFYTLARFGHDIANL
LGRLSENAQQDTKIVTAIDELQRLKAHLQVQWQCAVSIDVTELSGYHYHTGIVFNGY
INSETQPLVRGGRFDGMKSNQLATNQPRQATGFSMDVSRLLAHTQLDAPFIVLIDYD
AFNNLDSAQRQLLLQQVASLRQQGYRVTMPLTAEDMPVGLTHRLSLADNQWRLHA
V

6H*Pa*HisG$_S$ (26831 Da, ε280 = 10430 M^{-1} cm^{-1})

MHHHHHHENLYFQ\GMTEVTNSLPTSGLLNEANDEFLGLTLALSKGRILEETMPLLR
AAGVELLEDPEASRKLIFPTSNPNVRVLILRASDVPTYVEHGAADFGVAGKDVLLEHG
ANHVYELLDLKIAQCKLMTAGVKDAPLPNRRLRIATKYVNVARAYFASQGQQVDVI
KLYGSMELAPLVGLGDLIVDVVDTGNTLRANGLEARDHICDVSSRLIVNQVSYKRKF
ALLEPILDSFKNSINSTS

*Pa*HisZ (43085 Da, ε280 = 25900 M^{-1} cm^{-1})

GMLPDGVADVLFEDAHKQEVLRHQLTQQLITHGYQLVSPPMIEFTESLLSGASEDLK
RQTFKIIDQLTGRLMGIRADITPQILRIDAHHGGDGIARYCYAGDVIHTLPSGLFGSRT
PLQLGAEIFGCESIAADIELIDVLFSMINSLDMSAVLHVDLGHVTIFKRLAELAALSASD
TEQLMQLYANKNLPELKQVCQVLPMGSDFYTLARFGHDIANLLGRLSENAQQDTKI
VTAIDELQRLKAHLQVQWQCAVSIDVTELSGYHYHTGIVFNGYINSETQPLVRGGRF
DGMKSNQLATNQPRQATGFSMDVSRLLAHTQLDAPFIVLIDYDAFNNLDSAQRQLL
LQQVASLRQQGYRVTMPLTAEDMPVGLTHRLSLADNQWRLHAV

*Pa*HisG$_S$ (25213 Da, ε280 = 8940 M^{-1} cm^{-1})
GMTEVTNSLPTSGLLNEANDEFLGLTLALSKGRILEETMPLLRAAGVELLEDPEASRKL
IFPTSNPNVRVLILRASDVPTYVEHGAADFGVAGKDVLLEHGANHVYELLDLKIAQC
KLMTAGVKDAPLPNRRLRIATKYVNVARAYFASQGQQVDVIKLYGSMELAPLVGLG
DLIVDVVDTGNTLRANGLEARDHICDVSSRLIVNQVSYKRKFALLEPILDSFKNSINSTS

2.2 General Purification and Plasmid Engineering Methods

2.2.1 Starter Culture

LB medium (5–10 mL) containing either Ampicillin (100 μg/mL) or Kanamycin (50 μg/mL) was inoculated with a colony grown in LB agar, also containing Ampicillin (100 μg/mL) or Kanamycin (50 μg/mL), or from a glycerol stock (made up of 500 μL of a previous starter culture and 500 μL of 50% (v/v) glycerol solution and stored at −80 °C) and incubated in a shaker for 16 hours at 37 °C. The antibiotics used with each plasmid are shown in Table 2.1.

Table 2.1: Corresponding antibiotic for expression plasmids and strains

Plasmid/Strain	Antibiotic	Conc. (mg/mL)
pJexpress411 plasmid	Kanamycin	50
pJexpress414 plasmid	Ampicillin	100
pJexpress431 plasmid	Kanamycin	50
Origami 2 cells	Tetracycline	12.5

2.2.2 Culture Growth

LB medium (either 50 mL for expression test or 1 L for protein purification) containing antibiotic at the same concentration as starter culture was inoculated with starter culture (1–5 mL) and incubated in a shaker at 37 °C until the optical density at 600 nm (OD_{600}) of the culture reached 0.4–0.6 for expression at 37 °C, or OD 0.6–0.8 for expression at 15–20 °C. Cultures were equilibrated to the desired expression induction temperature, if other than 37 °C, prior to addition of 0.11 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and incubated in a shaker for 16 hours at the induction temperature. Cells were harvested by centrifugation (6800 g, 15 min), the supernatant was discarded, and pellets stored at −20 °C.

2.2.3 Expression Tests

One aliquot (1 mL) per culture was removed just before induction, centrifuged (~16,000 g, 1 min), supernatant discarded and pellet stored at −20 °C (control sample). This procedure was also carried out with post-induction culture (expression sample).

Cell pellets were resuspended in 100 μL of lysis buffer (Section 2.3), agitated gently for

20 min, centrifuged (16,000 g, 30 min) and the supernatant collected. The resulting cell pellets were resuspended either in 50 μL (control sample) or 100 μL (expression sample) of 10 mM Tris HCl pH 7.5. Samples were analysed by SDS-PAGE (either with the Invitrogen NuPage system or the Bio-Rad Criterion system).

2.2.4 Buffers for Protein Purification

Buffer A1: 50 mM HEPES, 500 mM NaCl, 10 mM Imidazole, pH 8.
Buffer A2: 20 mM HEPES, 150 mM NaCl, pH 8.
Buffer B: 50 mM HEPES, 500 mM NaCl, 500 mM Imidazole, pH 8.
Dialysis Buffer 1: 20 mM HEPES, 150 mM NaCl, 2 mM DTT, 10 % Glycerol (v/v), pH 7.5.
Dialysis Buffer 2: 20 mM HEPES, 150 mM NaCl, 10 % Glycerol (v/v), pH 8.0.

2.2.5 Protein Purification of 6xHis-tagged Proteins

All procedures were carried out at 4 °C. The protein of interest was purified from the cell pellet by thawing on ice and resuspending in Buffer A (20 mL per cell pellet from 1 L of culture) containing EDTA-free protease inhibitor cocktail (Cϕmplete, half tablet). Lysozyme (0.2 mg/mL) and DNase I (0.05 mg/mL) were added to the suspension and stirred for 30 min. The cells were disrupted by passage through a high-pressure cell disruptor (emulsifier) before centrifugation (48000 g, 30 min). The supernatant was collected, filtered in a 0.45 μm membrane and an aliquot (10 μL) stored (4 °C) as a pre-column sample. The sample was loaded onto a HisTrap FF 5 mL column, washed with 10 CV of Buffer A and eluted with a gradient of 20 CV of Buffer B (0–60%).

2.2.6 Protein Purification with TEVP 6xHis-tag Cleavage

Protein purifications incorporating TEVP cleavage of 6xHis-tagged proteins were initially purified as in Section 2.2.5. Fractions containing the desired protein were pooled and concentrated using a 10000 or 30000 molecular weight cut-off (MWCO) ultrafiltration membrane. The sample was dialysed overnight in 10000 or 30000 MWCO dialysis cassettes, or dialysis tubing, against 2 × 2 L of Dialysis Buffer 1. TEVP from in-house purification [20] was added slowly over multiple days to a final ratio of ~10:1 tagged protein to TEVP. The cleavage was dialyzed against fresh 2 L of Dialysis Buffer 1 before every TEVP addition. After cleavage the protein was dialysed into 2 × 2 L of Buffer A then centrifuged (48,000g, 30 min), filtered through 0.45 μm membrane and loaded onto a HisTrap FF 20 mL column (GE) pre-equilibrated with Buffer A. The flow through was collected and analysed by SDS–PAGE. Fractions containing the desired protein were pooled and concentrated using a 10000 or 30000 MWCO ultrafiltration membrane before

being dialyzed 3 × against 2 L of 20 mM HEPES (pH 8.0). The protein was aliquoted and stored at 80 °C.

2.2.7 Agarose Gel Electrophoresis Procedure

For agarose gel electrophoresis, 1% agarose gel (m/v) was prepared using TBE (Tris-borate-EDTA) buffer containing 0.06% Peqgreen stain (v/v). Gels were run at 80 V for 1 hour and visualised under UV light.

2.3 Expression and Purification Trials

2.3.1 *Sa*HisZ Expression Test Growth Condition Trials

The expression test conditions outlined in Table 2.2 were tried for the *Sa*HisZ.

Table 2.2: *Sa*HisZ Expression Test Growth Conditions

E. coli Strain	Plasmid	Media	Temp. (°C)	[IPTG] (mM)	Lysis Buffer
BL21(DE3)	SaHisZ-pJexpress414	LB	37	1	Bug Buster
BL21(DE3)	SaHisZ-pJexpress414	LB	16	0.5	Bug Buster

*Sa*HisZ Purification Trial

*Sa*HisZ was purified using the general purification method for a 6xHis-tagged protein requiring cleavage by TEVP (Section 2.2.6) with the following changes. After the first nickel column and ultrafiltration step the protein solution was centrifuged (48,000 g, 30 min) before dialysis (in dialysis tubing) into Dialysis Buffer 1 and addition of TEVP (7.6 mg). After two days cleaving with TEVP the protein solution was dialysed into Dialysis Buffer 2 for 2 hours then centrifuged (48,000 g, 30 min). Soluble and insoluble fractions were analysed by SDS-PAGE then cleavage by additional TEVP (3.8 mg). The purification was continued with the soluble fraction dialysed into 2 L Buffer A2. A further aliquot of TEVP (1.9 mg) was added and left to dialyse in fresh Buffer A2 for 2 days. The solution was filtered and run on a nickel column and analysed by SDS-PAGE. All fractions containing *Sa*HisZ were pooled and re-run on a third nickel column as in Section 2.2.6. Finally, the protein solution was loaded onto a HiPrep 26/60 Sephacryl S300 HR gel filtration column with 20 mM HEPES (pH 7.5) and analysed by SDS-PAGE. Fractions containing *Sa*HisZ were dialysed and stored as in Section 2.2.6. Bands from

the final SDS-PAGE were excised and sent for protein identification and intact mass determination and by mass spectrometry and protein identification by gel band digestion with trypsin, electrospray ionisation mass spectrometry (ESI-MS) and a Mascot search with comparison to the NCBI database.

2.3.2 *Sa*HisG$_S$ Expression Test Growth Condition Trials

The expression test conditions outlined in Table 2.3 were tried for the given protein.
All strains trialled contained a λDE3 lysogen carrying the gene for T7 RNA polymerase under control of the lacUV5 promoter. C43(DE3) is a strain specifically engineered for recombinant proteins which may be toxic to BL21(DE3) cells (widely used T7 *E. coli* expression strain) [112]. Origami 2(DE3) cells is a cell line with additional folding apparatus [113]. BL21Gold(DE3) cells are a superior strain of BL21(DE3) with two proteases capable of degrading recombinant protein removed [114]. BL21GoldPlysS(DE3) is similar but has shown to be advantageous for leaky expression of toxic proteins [114].

Auto-Induction Method of Protein Expression

Stock solutions of 20x NPS solution (0.5 M $(NH_4)_2SO_4$, 1 M KH_2PO_4, 1 M Na_2HPO_4) and 50x 5052 solution (2.71 M glycerol, 139 mM glucose, 292 mM α-lactose) were prepared.

Overnight culture (2 ml) of *Sa*HisG-pJ411 Bl21(DE3) was added to 400 ml Auto-induction Media (1 mM $MgSO_4$, 1 x 5052 solution, 1x NPS solution, 50 g/ml Kanamycin, ZY media to 400 ml) and agitated at 250 rpm and 16 °C for 48 hours. An aliquot (1 ml) was removed from each flask and processed with Bug Buster using the general expression test procedure (Section 2.2.3). The remaining volume of culture was centrifuged (6800 g, 15 min), the supernatant discarded and the cell pellet resuspended in Buffer A (20 ml). The samples were passed twice through an emulsifier, centrifuged (4800 g, 30 min) and the supernatant decanted and stored. A small scraping of the cell pellet was resuspended in Tris HCl (10 mM, pH 7.5, 200 ml) for SDS-PAGE. SDS-PAGE was performed for both sets of lysis conditions. This was repeated for *Sa*HisZ-pJ414 BL21(DE3).

2.3.3 *Sa*HisG$_S$-pJexpress411 Plasmid Engineering

The sequence encoding for Maltose-binding protein (MBP) was inserted into the *Sa*HisG$_S$-pJexpress411 plasmid to aid solubility of *Sa*HisG$_S$.

Table 2.3: Expression test growth conditions for *Sa*HisG$_S$ using *Sa*HisG$_S$-pJexpress411

E.coli Strain	Media	Temp. (°C)	[IPTG] (mM)	Lysis Buffer
BL21(DE3)	LB	37	1	Bug Buster
BL21(DE3)	LB	16	1	Bug Buster
BL21(DE3)	Auto-ind.	16	Auto-ind.	Bug Buster
BL21(DE3)	Auto-ind.	16	Auto-ind.	Emulsifier
BL21(DE3)	LB	37	0.5	Bug Buster
BL21(DE3)	LB	16	0.5	Bug Buster
BL21(DE3)	LB	37	0.1	Bug Buster
BL21(DE3)	LB	16	0.1	Bug Buster
C43 (DE3)	LB	37	0.5	Bug Buster
C43 (DE3)	LB	16	0.1	Bug Buster
C43 (DE3)	LB	37	0.5	Bug Buster
C43 (DE3)	LB	16	0.1	Bug Buster
Origami 2(DE3)	LB	37	0.5	Bug Buster
Origami 2 (DE3)	LB	16	0.1	Bug Buster
Origami 2 (DE3)	LB	37	0.5	Bug Buster
Origami 2 (DE3)	LB	16	0.1	Bug Buster
SoluBL21(DE3)	LB	37	0.5	Bug Buster
SoluBL21(DE3)	LB	16	0.5	Bug Buster
BL21Gold(DE3)	LB	37	0.5	Bug Buster
BL21Gold(DE3)	LB	37	0.1	Bug Buster
BL21Gold(DE3)	LB	16	0.5	Bug Buster
BL21Gold(DE3)	LB	16	0.1	Bug Buster
BL21Gold(DE3)pLysS	LB	37	0.5	Bug Buster
BL21Gold(DE3)pLysS	LB	16	0.5	Bug Buster
C43 (DE3)	TB	37	0.5	Bug Buster
C43 (DE3)	TB	16	0.5	Bug Buster
C43 (DE3)	TPB	37	0.5	Bug Buster
C43 (DE3)	TPB	16	0.5	Bug Buster
C43 (DE3)	LB & Mg^{2+}	37	0.5	Bug Buster

PCR to Linearize Parent Plasmid

The samples in Table 2.4 were prepared in written order from the following stock solutions: dNTP's (10 mM), Phusion HF Buffer (5x, Thermo Fisher Scientific), Primer 2 F (10 μM), Primer 1 R (10 μM), *Sa*HisG$_S$-pJexpress411 (20 ng/μL), Phusion polymerase (2 U/μL). Polymerase chain reaction (PCR) experiments of the samples were carried out using the thermocycler program detailed in Table 2.5. The samples were stored at −20 °C overnight before agarose gel electrophoresis was carried out. A Gene Ruler 1 KB DNA ladder was used to determine the size of resulting fragments.

Table 2.4: PCR to linearize *Sa*HisG$_S$-pJexpress411 sample composition. C1 was a control with plasmid absent and C2 was a control with polymerase absent.

Component	PCR 1	PCR 2	PCR 3	PCR 4	PCR 5	PCR 6	C1	C2
Annealing Temp. (°C)	52	53.2	55	52	53.2	55	55	55
Sterile Water (μL)	68	68	68	65	65	65	69	69
dNTP's (μL)	2	2	2	2	2	2	2	2
Phusion HF Buffer (μL)	20	20	20	20	20	20	20	20
Primer 2F (Tm = 52.9 °C) (μL)	5	5	5	5	5	5	5	5
Primer 1R (Tm = 51.7 °C) (μL)	5	5	5	5	5	5	5	5
DMSO (μL)	0	0	0	3	3	3	0	0
*Sa*HisG$_S$-pJexpress411 (μL)	1	1	1	1	1	1	0	1
Phusion Polymerase (μL)	1	1	1	1	1	1	1	0

Table 2.5: PCR conditions for *Sa*HisG$_S$-pJexpress411 linearization

Step	Temperature (°C)	Time (s)	Number of Cycles
Initial Denaturation	98	30	1
Denaturation	98	10	30
Annealing	(Table 2.4)	30	
Extension	72	30	
Final Extension	72	600	1
Hold	4	N/A	N/A

Gibson Assembly of Linear *Sa*HisG$_S$-pJexpress411 with MBP G-Block

PCR sample 6 (55 °C with DMSO, Section 2.3.3), was digested overnight with Tango buffer (10.5 μL) and DpnI (5 μL) at 37 °C. Two samples were prepared for Gibson Assembly as in Table 2.6.

Table 2.6: Gibson Assembly sample compositions

Component	5x	10x	Control
Sterile Water (μL)	6	4	10
*Sa*HisG$_S$-MBP G Block (μL)	2	4	0
Linear *Sa*HisG$_S$-pJexpress411 (μL)	2	2	0
Gibson Master Mix (μL)	10	10	10

The solutions were incubated in a thermocycler at 50 °C for 1 hour. The solution was mixed with NEB 5-α chemically competent cells (50 μL) on ice for 30 min then heat shocked for 32 s at 42 °C. SOC medium (300 μL) was added and the culture incubated at 37 °C for 1 hour. The solutions were plated on LB agar plates containing 50 mg/mL Kanamycin and incubated at 37 °C overnight.

Each colony was labelled with a letter (A–Z and α) and grown in LB media containing 50 mg/mL Kanamycin (5 mL) at 37 °C overnight. To pellet the cells, 5 mL was centrifuged at 16000 g for 1 min and the supernatant discarded.

Plasmid Purification of Gibson Assembly DNA

Two healthy looking colonies, M and X, from successful Gibson Assembly colonies (Section 2.3.3 were purified with a Qiagen Mini-Prep Purification Kit as per manufacturer instructions. The samples in Table 2.7 were incubated for 1 hour at 37 °C before an agarose gel was run and visualised under UV light.

Purified plasmids *Sa*HisG$_S$-pJexpress411-MBP-(construct X) and *Sa*HisG$_S$-pJexpress411-MBP-(construct M) from successful Gibson Assembly were sent to Eurofins Genomics for sequencing with the T7 promoter primer (5′-TAATACGACTCACTATAGGG-3′) and an internal sequencing primer (5′-CAT TTA AAG GCC AGC CTA GCA-3′).

Table 2.7: Gibson Assembly plasmid purification incubation samples. C is a control without Nde1. D1 and D2 represent digestions at two concentrations of plasmid.

Incubation Condition	Plasmid X C	Plasmid X D1	Plasmid X D2	Plasmid M C	Plasmid M D1	Plasmid M D2
Plasmid X (μL)	5	5	10	0	0	0
Plasmid M (μL)	0	0	0	5	5	10
HindIII (μL)	0	1	1	0	1	1
Nde1 (μL)	0	1	1	0	1	1
Buffer (μL)	2	2	2	2	2	2
Water (μL)	13	11	6	13	11	6

Expression Tests of *Sa*HisG$_S$-pJexpress411-MBP

The expression trials were carried out as outlined in Table 2.8.

Table 2.8: Expression tests of *Sa*HisG$_S$-pJexpress411-MBP.

E.coli Strain	Media	Temp. (°C)	[ITPG] (mM)	Induction Time (Hours)	Lysis Buffer
BL21(DE3)	LB	37	0.5	20	Bug Buster
BL21(DE3)	LB	37	0.5	44	Bug Buster
BL21(DE3)	LB	20	0.5	20	Bug Buster
BL21(DE3)	LB	15	0.5	20	Bug Buster
Origami 2	LB	37	0.5	20	Bug Buster
Origami 2	LB	16	0.5	20	Bug Buster
SoluBL21	LB	37	0.5	20	Bug Buster
SoluBL21	LB	37	0.5	44	Bug Buster
SoluBL21	LB	16	0.5	20	Bug Buster
SoluBL21	LB	16	0.5	44	Bug Buster
C43(DE3)	LB	16	0.5	20	Bug Buster

Protein Purification and Solubility Trials

Large scale trials (1–2 L of culture) were passed through a HisTrap FF 5 mL nickel column to determine whether soluble *Sa*HisG$_S$ was present in concentrations lower than what could be discernible by analysis of the crude extract. The conditions listed in Table 2.9 were tested.

Table 2.9: Large scale protein purification solubility trials with *Sa*HisG$_S$-pJexpress411.

E. coli Strain	Media	Induction Temp. (°C)	Culture Volume (L)	IPTG (mM)
C43(DE3)	LB	16	1	0.5
C43(DE3)	LB	37	2	0.5

2.3.4 Lysis Buffer Solubility Trials

Tris/Bis Tris Lysis Buffer Screen

Alternative lysis buffers to Bug Buster were screened to determine whether *Sa*HisG$_S$ requires a different reagent for solubility. The expression tests outlined in Table 2.10 were trialled.

Reducing Agent Buffer Trials

β-mercaptoethanol (BMCE) and dithiothreitol (DTT) were tested with original protein purification buffer (buffer A1) to determine whether they had a stabilising effect on soluble protein by reducing disulphide bonds. Cultures of *Sa*HisG$_S$-pJexpress411 grown at 37 °C in LB media and induced with 0.5 mM IPTG were used and the lysis buffers outlined in Table 2.11 were tested. HEPES buffers were adjusted to pH 8.0, Tris buffers to pH 8.0 and Bis-Tris buffers to pH 6.8.

*Sa*HisG$_S$ Protein Purification Trial with DTT in Buffer

Large scale trial (4 L of C43(DE3) culture) was passed through a HisTrap Excel FF 5 mL nickel column using buffers containing DTT to stabilise soluble *Sa*HisG$_S$. Line A contained 50 mM HEPES, 150 mM NaCl, 10 mM Imidazole, 4mM DTT and Line B contained 50 mM HEPES, 150 mM NaCl, 500 mM Imidazole, 4 mM DTT.

*Sa*HisG$_S$ Lysis Buffer Trial with *Sa*HisZ in Buffer

*Sa*HisZ was tested with the original protein purification buffer to determine whether *Sa*HisZ has a stabilising effect on soluble protein. Expression tests of *Sa*HisG$_S$-pJexpress411 grown at 37 °C in LB media and induced with 0.5 mM IPTG were used and the lysis buffers outlined in Table 2.12 were tested (in addition, all contained 150 mM NaCl).

Table 2.10: Tris and Bis-Tris lysis buffer expression tests with *Sa*HisG$_S$-pJexpress411. with *Sa*HisG$_S$-pJexpress411.

E. coli Strain	Medium	Expression Temp °C	[ITPG] (mM)	Lysis Buffer
C43 (DE3)	LB	37	0.5	20 mM Bis Tris, 150 mM NaCl
C43 (DE3)	LB	37	0.5	20 mM Bis Tris, 500 mM NaCl
C43 (DE3)	LB	37	0.5	20 mM Bis Tris, 150 mM NaCl, 100 mM Arginine
C43 (DE3)	LB	37	0.5	20 mM Bis Tris, 150 mM NaCl, 100 mM Glycine
C43 (DE3)	LB	37	0.5	20 mM Bis Tris, 150 mM NaCl, 0.1% v/v Igepal
C43 (DE3)	LB	37	0.5	20 mM Bis Tris, 150 mM NaCl, 0.1% v/v Triton X100
C43 (DE3)	LB	37	0.5	20 mM Bis Tris, 150 mM NaCl, 10 mM β-mercaptoethanol
C43 (DE3)	LB	37	0.5	20 mM Tris, 150 mM NaCl
C43 (DE3)	LB	37	0.5	20 mM Tris, 500 mM NaCl
C43 (DE3)	LB	37	0.5	20 mM Tris, 150 mM NaCl, 100 mM Arginine
C43 (DE3)	LB	37	0.5	20 mM Tris, 150 mM NaCl, 100 mM Glycine
C43 (DE3)	LB	37	0.5	20 mM Tris, 150 mM NaCl, 0.1% v/v Igepal
C43 (DE3)	LB	37	0.5	20 mM Tris, 150 mM NaCl, 0.1% v/v Triton X100
C43 (DE3)	LB	37	0.5	20 mM Tris, 150 mM NaCl, 10 mM β-mercaptoethanol
C43 (DE3)	LB	37	0.5	20 mM Tris, 150 mM NaCl, 8 μM *Sa*HisZ

Table 2.11: Reducing agent buffers expression tests.

Reducing Agent Buffer	HEPES (mM)	[NaCl] (mM)	BMCE (mM)	DTT (mM)	Imidazole (mM)	Histidine (mM)
1	50	-	10	-	-	-
2	50	500	10	-	10	-
3	50	500	-	4	10	-
4	50	500	-	-	10	-
5	50	-	-	4	-	-
6	50	500	10	-	-	-
7	50	150	10	-	10	-
8	75	-	-	6	-	-
9	75	225	-	6	-	-
10	75	-	-	6	15	-
11	75	225	-	6	15	-
12	75	-	-	6	-	15
13	75	-	-	6	-	15

Table 2.12: *Sa*HisZ buffers expression tests

*Sa*HisZ Buffer	Tris (mM)	HEPES (mM)	*Sa*HisZ (μM)	DTT (mM)	Imidazole (mM)	Histidine (mM)
1	20	-	8	-	-	-
2	20	-	8	-	10	-
3	-	50	8	-	-	-
4	-	50	8	-	10	-
5	-	50	8	4	-	-
6	-	50	8	-	-	10

Co-purification of *Sa*HisG$_S$ with *Sa*HisZ (inc. Glass beads)

Three protein purifications utilising co-purification with *Sa*HisZ were carried out as in Section 2.2.5 but with 50:50 *Sa*HisZ and *Sa*HisG$_S$ cell pellets used. Two trials with low salt were trialled, either with 150 mM NaCl or 150 mM KCl in Buffers A and B instead of 500 mM NaCl. An additional trial using glass beads to lyse the cells instead of an emulsifier was undertaken. For glass bead lysis, cells pellets were prepared in the same way as when using the emulsifier, but instead of passing through the cell disrupter, the cell pellets were shaken overnight at 4 °C with 1 ml of 0.5 mm diameter small glass beads. Due to the longer lysis period, one whole tablet of EDTA free protease inhibitor cocktail (Cϕmplete) was used along with a reduction in lysozyme (0.01 mg/ml).

2.4 Expression of *Pa*HisG$_S$ and *Pa*HisZ

*Pa*HisG$_S$-pJ411express plasmid was transformed into *E. coli* C43(DE3) competent cells (Sigma-Aldrich) and *Pa*HisZ-pJ434express plasmid was transformed into *E. coli* BL21(DE3) competent cells (Novagen). Transformed cells were grown and induced as in Section 2.2.2 at 16 °C.

*Pa*HisG$_S$ and *Pa*HisZ (tagged and un-tagged) were purified separately from respective cell pellets with all steps carried out at 4 °C and utilising an AKTA Start FPLC system (GE Lifesciences) for the chromatographic procedures. The final concentrations of *Pa*HisG$_S$ and *Pa*HisZ were determined spectrophotometrically using a NanoDrop at 280 nm with theoretical extinction coefficients (ε280) of 8940 M^{-1} cm^{-1} and 25900 M^{-1} cm^{-1}, respectively for untagged protein and 10430 M^{-1} cm^{-1} and 27390 M^{-1} cm^{-1}, respectively for tagged protein. The molecular masses of 25213 Da (*PaHisG$_S$*) and 43085 Da (*Pa*HisZ) were verified by ESI-MS.

2.4.1 Purification of *Pa*HisG$_S$ and *Pa*HisZ

The purification of *Pa*HisG$_S$ and *Pa*HisZ were carried out exactly as in Stroek et al. [20]. A HisTrap FF 5 mL column with a linear 0–300 mM imidazole gradient was used to remove contaminant proteins by fractionation before and after cleavage of the 6xHis-tag by TEVP. Fractions were analysed by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) (NuPAGE Bis-Tris 4–12% Precast gels, Thermo Fisher Scientific) and concentrated using 10000 molecular weight cut off (MWCO) ultrafiltration membranes (Millipore). Both proteins were stored in 20 mM HEPES (pH 8.0) at −80 °C.

2.4.2 Method Development of Large Scale *Pa*HisZ Purification

A larger scale purification of *Pa*HisZ was trialled with cell pellets from 5 L of cell culture and a larger HisTrap FF 20 mL column. Cells were thawed on ice for 30 min before being resuspended in Buffer A [50 mM HEPES, 10 mM imidazole, and 500 mM NaCl (pH 8.0)] containing 0.2 mg/mL lysozyme, 0.05 mg/mL DNase I, and one tablet of EDTA-free Cømplete protease inhibitor cocktail. Cells were disrupted in a high-pressure cell disruptor and centrifuged at 48000 g for 30 min to remove cell debris. The supernatant was filtered through a 0.45 μm membrane and loaded onto a HisTrap FF 20 mL column pre-equilibrated with Buffer A. The column was washed with 15 CV of Buffer A and the adsorbed proteins were eluted with 20 CV of a linear gradient from 0–40% Buffer B [50 mM HEPES, 500 mM imidazole, and 500 mM NaCl (pH 8.0)]. Fractions containing the desired protein were initially pooled and concentrated using repeated cycles of a 10000 MWCO centrifugal filter unit (Amicon) and centrifuge (5000 g, 15 min) before using a 30000 MWCO ultrafiltration membrane. The sample was dialysed overnight in 30000 MWCO dialysis cassettes against 2 times 2 L of 20 mM HEPES (pH 8.0) followed by 2 L of 20 mM HEPES (pH 7.5) before adding 10 mg of TEVP and dialyzed twice against 2 L of Dialysis Buffer 1. This was repeated for two further additions 10 mg of TEVP and then dialysed against 2 L of Buffer A. Sample was centrifuged (48000 g, 30 min) then filtered through 0.45 μm membranes and loaded onto a HisTrap FF 20 mL column pre-equilibrated with buffer A. The flow through was collected and analyzed by SDS–PAGE, concentrated using 30000 MWCO ultrafiltration membranes, dialyzed 3 times against 2 L of 20 mM HEPES (pH 8.0). A second round of TEVP cleavage was carried out with 7 mg TEVP for 5 days before dialyzing twice against 2 L of Buffer A. Sample filtered once more through 0.45 μm membranes and loaded onto a HisTrap FF 20 mL column pre-equilibrated with Buffer A and using same method as previously. Fractions containing *Pa*HisZ were pooled, concentrated as previously, aliquoted and stored at −80 °C.

2.4.3 Purification of 6H*Pa*HisG$_S$ and 6H*Pa*HisZ

Cells were thawed on ice for 20 min before being resuspended in Buffer A containing 0.2 mg/mL lysozyme, 0.05 mg/mL DNase I, and one half-tablet of EDTA-free Cømplete protease inhibitor cocktail. Cells were disrupted in a high-pressure cell disruptor and centrifuged at 48000 g for 30 min to remove cell debris. The supernatants were filtered through a 0.45 μm membrane and loaded onto a HisTrap FF 5 mL column pre-equilibrated with Buffer A. The column was washed with 10 CV of Buffer A and the adsorbed proteins were eluted with 20 CV of a linear gradient from 0 to 60% Buffer B. Fractions containing the desired protein with no visible contaminant proteins were pooled and dialyzed twice against 2 L of 20 mM HEPES (pH 8.0) before being concentrated using a 10000 molecular MWCO or 30000 MWCO ultrafiltration

membrane (Millipore) for 6H*Pa*HisG$_S$ and 6H*Pa*HisZ respectively. The samples were aliquoted and stored at −80 °C.

2.5 Differential Scanning Fluorimetry-based Thermal Shift Assay

2.5.1 DSF-based Thermal Shift Assay of Histidine Binding to *Pa*HisZ

Differential scanning fluorimetry-based (DSF) thermal shift assay was performed as previously described [115] at λ_{ex} = 490 nm and λ_{em} = 610 nm in 96-well plates on a Stratagene Mx3005p instrument. The assay of histidine binding to *Pa*HisZ was performed from 25–65 °C in 1 °C/min increments in quintuplicate with 9.4 μM *Pa*HisZ in Kinetic Buffer [100 mM Tricine, 15 mM $MgCl_2$, 100 mM KCl, 4 mM DTT] and 0–8 mM histidine. Sypro Orange (5×) (Invitrogen) was added to all wells. The T_m of each sample was determined as previously described [115] with control curves (no enzyme) subtracted from curves containing enzyme. Equation 2.1 was used and the data analysed by the nonlinear regression function of SigmaPlot 13 (SPSS Inc.). F_U is fraction unfolded, T is the temperature in °C, T_m is the melting temperature, c is the slope of the transition region, and LL and UL are folded and unfolded baselines, respectively. The results were plotted against [histidine] to determine the K_D with respect to *Pa*HisZ using SigmaPlot 13's nonlinear regression function.

$$F_U = LL + \frac{UL - LL}{1 + e^{\frac{(T_m - T)}{c}}} \tag{2.1}$$

2.5.2 DSF-based Thermal Shift Assay of PRPP Binding to *Pa*HisG$_S$

Differential scanning fluorimetry-based thermal shift assay was performed as previously described [115]. The assay for PRPP binding to *Pa*HisG$_S$ was performed from 25–70 °C (with a 1 °C/min increase) in quintuplicate with 10 μM *Pa*HisG$_S$ in Kinetic Buffer [100 mM Tricine, 15 mM $MgCl_2$, 100 mM KCl, 4 mM DTT] and 0–2 mM PRPP. The fraction of protein unfolded at each temperature was determined as described in Section 2.5.1 using Equation 2.1 and compared between two different batches of *Pa*HisG$_S$.

2.6 Kinetic Methods

Unless otherwise stated all kinetic methods were carried out using 5x Kinetic Buffer Stock Solution (500 mM Tricine, 75 mM $MgCl_2$, 500 mM KCl, 20 mM DTT, 10 μM pyrophosphatase, pH 8.5) to give a final concentration of Kinetic Buffer of 100 mM Tricine, 15 mM $MgCl_2$, 100 mM KCl, 4 mM DTT and 2 μM pyrophosphatase (*Mt*PPase). An UV-2600 spectrophotometer outfitted with a CPS unit for temperature control was used unless stated otherwise. PRATP formation was monitored continuously at 290 nm (ϵ_{290} = 3600 M^{-1} cm^{-1} at pH 8.5) at 20 °C in 1-cm path length quartz cuvettes. Rate of absorbance change was transformed into rate of PRATP formation using Equation 2.2. The standard error was calculated for all readings using Equation 2.3 where n is the number of replicates recorded. Control reactions lacked either ATP, PRPP, *Pa*HisG$_S$, or *Pa*HisZ. Controls were also carried out to ensure that the rate did not depend on *Mt*PPase.

$$v\left(\mu M.s^{-1}\right) = \frac{\left(\frac{\Delta abs\ min^{-1}}{0.0036\ \mu M^{-1} cm^{-1}}\right)}{60\ s} \tag{2.2}$$

$$Standard\ Error = \frac{Standard\ Deviation}{\sqrt{n}} \tag{2.3}$$

2.6.1 Determination of K_D values

The dissociation constant (K_D) values for the *Pa*HisG$_S$ and *Pa*HisZ subunits of *Pa*ATPPRT were determined by using 0.67 μM HisG and monitoring the rate of PRATP formation from 0–18.4 μM *Pa*HisZ.

The $K_D$value for the *Pa*HisG$_S$ and *Pa*HisZ subunits of 6H*Pa*ATPPRT were determined by using 0.17 μM *Pa*HisG$_S$ and monitoring the rate of PRATP formation from 0–19 μM *Pa*HisZ.

The K_D was calculated by fitting to the Morrison Equation (Equation 2.4) [116] where E is equal to [*Pa*HisG$_S$], S is equal to [*Pa*HisZ]. And V_{max} is the maximal rate of the system.

$$v = V_{max}\frac{(E+S+K_D) - \sqrt{(E+S+K_D)^2 - 4ES}}{2E} \tag{2.4}$$

2.6.2 Determination of K_M values

The K_M values with respect to each substrate (ATP or ADP and PRPP) were determined by varying the concentration of one substrate while fixing the other substrate concentration at saturating levels (see Table 2.13).

Table 2.13: Assay concentrations of substrates used for determining K_M values.

K_M	Enzyme	[*Pa*ATPPRT] (μM)	[ATP] (mM)	[ADP] (mM)	[PRPP] (mM)
${K_M}^{ATP}$	*Pa*ATPPRT	0.28	0.4–5.6	-	2
${K_M}^{PRPP}$	*Pa*ATPPRT	0.28	5.6		0.1–2
${K_M}^{PRPP}$	*Pa*ATPPRT	0.28	-	5.6	0.1–4
${K_M}^{ADP}$	*Pa*ATPPRT	0.28	-	0.4–5.6	2
${K_M}^{ATP}$	6H*Pa*ATPPRT	0.16	0.2 - 1.2	-	2
${K_M}^{PRPP}$	6H*Pa*ATPPRT	0.16	5.6	-	0.03–4

The K_M was calculated using Equation 2.5 if no substrate inhibition was observed, or if it was, either Equation 2.6 (standard substrate inhibition) or Equation 2.7 (Haldane substrate inhibition [117]) where v is the initial rate, E is the concentration of enzyme, k_{cat} is the steady-state turnover number, S is the substrate concentration, K_M is the Michaelis constant and K_i is the inhibitor dissociation constant.

$$\frac{v}{E} = \frac{k_{cat}S}{(K_M + S)} \tag{2.5}$$

$$\frac{v}{E} = \frac{k_{cat}S}{(K_M + S)\left(1 + \frac{S}{K_i}\right)} \tag{2.6}$$

$$\frac{v}{E} = \frac{k_{cat}S}{K_M + S + \frac{S^2}{K_i}} \tag{2.7}$$

2.6.3 Determination of IC_{50} values

The inhibitor IC_{50} values for histidine and AMP with respect to *Pa*ATPPRT were determined by monitoring the rate of PRATP formation with varying concentrations of each inhibitor and fixed enzyme and substrate concentrations (see Table 2.14).

Table 2.14: Assay concentration of inhibitors, enzyme and substrates used for determining IC_{50} values.

Enzyme	Inhibitor	[E] (μM)	[I] (μM)	[ATP] (mM)	[ADP] (mM)	[PRPP] (mM)
*Pa*ATPPRT	Histidine	0.24	0–320	5.6	-	2
*Pa*ATPPRT	Histidine	0.28	0–320	-	5.6	2
6H*Pa*ATPPRT	Histidine	0.16	0–640	5.6	-	2
*Pa*ATPPRT	AMP	0.14	0–2400	5.6	-	2
*Pa*HisG$_S$	AMP	4.5	0–800	5.6	-	2
*Pa*HisG$_S$	ITP	4.5	0–800	5.6	-	2

The IC_{50} value is calculated using Equation 2.8 where v_o and v_i are the rates in the absence and presence of inhibitor, respectively, I is the concentration of inhibitor and IC_{50} is the half-maximal inhibitory concentration.

$$\frac{v_i}{v_o} = \frac{1}{1+\left(\frac{I}{IC_{50}}\right)^h} \tag{2.8}$$

2.6.4 Determination of $K_i$values

Determination of K_i values for histidine and AMP were carried out by monitoring the rate of PRATP formation across a range of substrate concentrations at fixed inhibitor concentrations (see Table 2.15). K_i values were obtained by fitting globally to a 3D hyperbola for initial rate (v), I and S (inhibitor and substrate concentration, respectively). The resulting lines of best fit were fitted to one of Equation 2.9 (competitive inhibition), Equation 2.10 (non-competitive inhibition) or Equation 2.11 (uncompetitive inhibition) to determine which model for inhibition best fits te data. S is the substrate concentration, E is the enzyme concentration, K_{ii} is the K_i intercept and K_{is} is the K_i slope, k_{cat} is the steady-state turnover number, K_M is the Michaelis constant and K_i is the inhibitor dissociation constant. The data was then replotted in Lineweaver Burke form.

$$\frac{v}{E} = \frac{k_{cat}\ S}{\left(1+\left(\frac{I}{K_i}\right)\right)K_m + S} \tag{2.9}$$

$$\frac{v}{E} = \frac{k_{cat}\ S}{S\left(1+\left(\frac{I}{K_{ii}}\right)\right) + K_m\left(1+\left(\frac{I}{K_{is}}\right)\right)} \tag{2.10}$$

$$\frac{v}{E} = \frac{k_{cat}\ S}{S\left(1 + \left(\frac{I}{K_i}\right)\right) + \ K_m} \tag{2.11}$$

Table 2.15: Concentration of substrates, enzyme and inhibitors used for inhibition studies.

Enzyme	Substrate Studied	Inhibitor	[E] (μM)	[I] (μM)	[ATP] (mM)	[PRPP] (mM)
*Pa*ATPPRT	PRPP	Histidine	0.28	0–20	2.8	0.1–2
*Pa*ATPPRT	ATP	Histidine	0.28	0–20	0.4–5.6	2
*Pa*HisG$_S$	ATP	AMP	4.5	0–250	0.4–5.6	2
*Pa*HisG$_S$	PRPP	AMP	4.5	0–50	5.6	0.1–2
*Pa*ATPPRT	ATP	AMP	0.29	0–400	0.4–5.6	2
*Pa*ATPPRT	PRPP	AMP	0.29	0–200	5.6	0.1–2

2.6.5 *Pa*HisG$_S$ Mutant Sensitivity to Histidine

In order to determine whether *Pa*HisG$_S$ mutants C115S and R56A are as sensitive to histidine as the WT, the rate of PRATP formation was monitored using 4.7 μM *Pa*ATPPRT (either WT, C115S or R56A *Pa*HisG$_S$ complexed with WT *Pa*HisZ) with 0 μM or 640 μM histidine. The rate of PRATP formation by the mutants were compared to WT data and normalised against 0 μM histidine.

2.6.6 Effect of ITP, CTP and R5P on *Pa*HisG$_S$

To determine whether ITP (inosine triphosphate) influences *Pa*HisG$_S$, the rate of PRATP formation was monitored using 4.5 μM *Pa*HisG$_S$ with 0–6 mM ITP in the standard assay conditions of 5.6 mM ATP and 2 mM PRPP.

To determine whether CTP (cytidine triphosphate) influences *Pa*HisG$_S$, the rate of PRATP formation was monitored using 4.5 μM *Pa*HisG$_S$ with 0 or 2.4 mM CTP. Due to the absorbance of CTP, smaller path length (0.5 cm) cuvettes were used with 2 mM ATP and 4 mM PRPP.

To determine whether ribose-5-phosphate (R5P) influences *Pa*HisG$_S$, the rate of PRATP formation was monitored using 4.5 μM *Pa*HisG$_S$ with 0–2.4 mM R5P in the standard assay conditions of 5.6 mM ATP and 2 mM PRPP.

2.6.7 Effect of 3-(2-thienyl)-L-alanine on *Pa*ATPPRT

To determine whether 3-(2-thienyl)-L-alanine (TIH) influences *Pa*ATPPRT the rate of PRATP formation was monitored using 0.29 μM *Pa*ATPPRT with 0 or 4 mM TIH under the standard assay conditions.

2.6.8 Effect of dATP as an Alternative Substrate

To determine whether dATP is an alternative substrate for *Pa*HisG$_S$ the formation of PRATP was monitored for assays of 4.5 μM *Pa*HisG$_S$ with 0–5.6 mM dATP (instead of ATP) as the substrate and 2 mM PRPP.

2.6.9 Effect of buffer pH on *Pa*HisG$_S$ Activity

To determine whether *Pa*HisG$_S$ activity was effected by an unexpected error in pH preparation of the buffer, rate of absorbance change was monitored for an assay of 7.5 μM *Pa*HisG$_S$, 2 μM PRPP and 5.6 mM ATP with Kinetic Buffer [100 mM Tricine, 15 mM $MgCl_2$, 100 mM KCl, 4 mM DTT, 2 μM PPase] at pH 8.1–8.8. ϵ_{290} = 3100 M^{-1} cm^{-1} at pH 8.0, 3600 M^{-1} cm^{-1} at pH 8.5 and 4000 M^{-1} cm^{-1} at pH 9.5. This was repeated for a second, more active, batch, of *Pa*HisG$_S$ (2 μM).

2.6.10 Rapid kinetics of *Pa*HisG$_S$ and *Pa*ATPPRT Reaction with Histidine Present

Using an SX-20 stopped-flow spectrophotometer (0.9-ms deadtime), syringe A (2.4 μM - 4.2 μM *Pa*HisG$_S$ or 3.98 μM *Pa*ATPPRT, 0–1280 μM histidine, 4 mM PRPP, 8 mM DTT, 4 μM PPase, 100 mM Tricine, 15 mM $MgCl_2$, 100 mM KCl, pH 8.5) was rapidly mixed with syringe B (11.2 mM ATP, 100 mM Tricine, 15 mM $MgCl_2$, 100 mM KCl, pH 8.5) in a 0.5-cm mixing cell in quadruplicate traces. Final concentrations in the mixing cells were 1.2 μM - 2.1 μM *Pa*HisG$_S$ or 1.99 μM *Pa*ATPPRT, 0–640 μM histidine, 2 mM PRPP, 5.6 mM ATP, 4 mM DTT, 2 μM PPase, 100 mM Tricine, 15 mM $MgCl_2$, 100 mM KCl. The time course of PRATP formation was monitored for 10 s. The rate of PRATP formation with and without histidine was compared between *Pa*HisG$_S$ and *Pa*ATPPRT. A t-test was performed to confirm statistical significance.

2.6.11 Effect of Histidine on the Approach to Steady State

Using an SX-20 stopped-flow spectrophotometer, 55 μL of syringe A (4 μM *Pa*ATPPRT, 0 or 1280 μM histidine, 4 mM PRPP, 4 mM DTT, 10 μM *Mt*PPase, 100 mM Tricine, 15 mM $MgCl_2$, 100 mM KCl, pH 8.5) was rapidly mixed with 55 μL of syringe B (11.2 mM ATP, 4 mM DTT, 10 μM PPase, 100 mM Tricine, 75 mM $MgCl_2$, 100 mM KCl, pH 8.5) in a 5 μL mixing cell, in quadruplicate. Final concentrations in the mixing cells were 2 μM *Pa*ATPPRT, 0 or 640 μM histidine, 2 mM PRPP, 5.6 mM ATP, 4 mM DTT, 10 μM *Mt*PPase, 100 mM Tricine, 15 mM $MgCl_2$, 100 mM KCl. The time course of PRATP formation was monitored for 10 s. The rate of PRATP formation was monitored by absorbance increase in a linear-time base for 10 s with 1000 data points collected per trace. Rates with and without histidine were compared. Controls lacked PRPP.

2.6.12 Approach to Steady State in the Presence of Histidine

Using an SX-20 stopped flow spectrophotometer, 55 μL of syringe A (38 μM *Pa*ATPPRT, 0 or 60 μM histidine, 4 mM PRPP, 4 mM DTT, 10 μM *Mt*PPase, 100 mM Tricine, 15 mM $MgCl_2$, 100 mM KCl, pH 8.5) was rapidly mixed with 55 μL of syringe B (0 or 60 μM histidine, 7 mM ATP, 4 mM DTT, 10 μM PPase, 100 mM Tricine, 75 mM $MgCl_2$, 100 mM KCl, pH 8.5) in a 5 μL mixing cell, in quadruplicate. Final concentration in the mixing cell were 19 μM *Pa*ATPPRT, 0 or 30 μM histidine, 2 mM PRPP, 3.5 mM ATP, 4 mM DTT, 10 μM *Mt*PPase, 100 mM Tricine, 15 mM $MgCl_2$, 100 mM KCl. Controls lacked PRPP. The time course of PRATP formation was monitored by absorbance increase in a split-time base for 1 s, with 2500 data points collected before 0.2 s and 2500 data points in the following 0.8 s. The rate of PRATP formation with and without histidine was compared between *Pa*HisG$_S$ and *Pa*ATPPRT.

2.7 Crystallography

2.7.1 Crystallisation Trials of *Pa*ATPPRT:Histidine

Crystallisation trials for histidine bound to the hetero-octameric complex were based upon previously successful crystallization trials of *Pa*ATPPRT [20]. All steps were carried out at 4 °C. *Pa*HisG$_S$ and *Pa*HisZ were mixed in 1:1 molar ratio and incubated with 10 mM Mg-ATP overnight. The buffer was exchanged with Crystallisation Buffer [20 mM Tris (pH 7.0), 50 mM KCl, 10 mM $MgCl_2$, 2 mM DTT, 0.5 mM histidine] and concentrated using an 10,000 MWCO centrifugal filter unit to $\sim$ 8 mg/ml before 0.5 μL

Mg-ATP per 50 μL of protein was supplemented in.

Crystal trials were set up using the hanging-drop vapour-diffusion method in 24-well plates with 500 μL Reservoir Buffer [11% polyethylene glycol (PEG) 3350, 100 mM bicine (pH 8.5), 150 mM $SrCl_2$, 150 mM KBr, 2% 1,6-hexanediol] in the reservoir. Drops were comprised of 1 μL protein solution and 1 μL Reservoir Buffer with or without 0.2 μL 1:500 seed stock. Crystals were left to grow for 21 days at 4 °C.

2.7.2 Soaking Trials of Substrate and Histidine into *Pa*ATPPRT Crystals

Three crystals were washed in Soaking Buffer 1 [10% PEG 3350, 0.1 M bicine pH 8.5, 50 mM MgCl2, 0.1 M KBr, and 4% 1,6-hexanediol] before being transferred to a series of droplets containing soaking buffer, substrate and histidine. Solid particles of ATP, PRPP and histidine were added to Soaking Buffer 1 to create Soaking Buffer 2 [ATP and PRPP] and Soaking Buffer 3 [ATP, PRPP and histidine].

Crystals were soaked in Soaking Buffer 1 for 1 min, which was then followed by soaking in Buffers 2 and 3 for 3, 30 or 60 min each. A fourth trial was set up where the crystal was soaked for 10 min in Soaking Buffer 2 and 20 s in Soaking Buffer 3. Soaking in Buffers 1-3 was always followed by 10 s soak in cryoprotectant [20% 2-methyl-2,4-pentanediol (MPD), 80 % Soaking Buffer 1 (v/v)] and then flash frozen in liquid nitrogen at -196 °C.

2.7.3 X-ray Data Collection and Processing

Crystals were tested in-house for X-ray diffraction and the data were collected on a Rigaku 007 MM HFM generator with a Saturn 944+ CCD detector. Crystals that diffracted less than 5 Å in-house were sent for data collection at beamline i24 at the Diamond Light Source (Oxfordshire, UK). The diffraction data was processed using the automated processing pipeline integrated with XDS [118] and scaled using AIMLESS [119]. The structure of good quality diffraction data were solved with MOLREP [120] using PDB 6FTT [101] as a search model. The model was refined using cycles of model-building with COOT [121] and refinement with REFMAC [122].

Chapter 3

RESULTS AND DISCUSSION

3.1 Expression and Purification of *Sa*ATPPRT

Initially the ATPPRT from *S. aureus* was sought for study. In addition to the rational presented in Section 1.2 for studying ATPPRTs (advancement in biotechnology and as a model for allostery and *A. baumannii* drug discovery), study of *S. aureus* ATPPRT provided an additional advantage. *S. aureus* is one of the leading bacterial pathogens in hospital-acquired infections [123], with around half of these strains being resistant to commonly prescribed antibiotics [124] and some resistant to all available antibiotics [125]. As such, new antibiotics for *S. aureus* are desperately needed. It was originally hoped that study of *S. aureus* ATPPRT in particular, could directly lead to a viable drug target as the histidine pathway is not present in humans and so side effects would be minimized. Although there were no knock-out studies published for *S. aureus* proving the gene essential to infection or virulence, there was evidence from flux balance analysis by Henriksen et.al [126] stating that inhibition of any histidine pathway step would halt histidine production. Therefore, purification of *Sa*ATPPRT was trialled first.

3.1.1 *Sa*HisZ Expression and Purification

Over-expression of *S. aureus* protein *Sa*HisZ (33460 Da with 6xHis-tag) occurred at both temperatures trialled (Figure 3.1), 37 °C and 16 °C. The protein mainly expressed in the insoluble fraction at 37 °C (Figure 3.1, lane 12). There was a small band present at the correct height for *Sa*HisZ in the soluble fraction (Figure 3.1, lane 11) but it is unclear whether this was a small expression of *Sa*HisZ or a contaminant protein. However, there is clear over-expression of *Sa*HisZ in the soluble fraction when induced at 16 °C with 0.5 mM IPTG Figure 3.1, lane 15). If protein is present in the soluble fraction it suggests that the protein has been folded correctly by the expression system and further

purification of the protein is possible. Lowering the temperature and the concentration of IPTG can often lead to increased soluble expression of protein because there is more time for protein folding due to slower expression. Additionally, high concentrations of IPTG can significantly reduce growth of recombinant *E. coli* cells and hence protein yield [127].

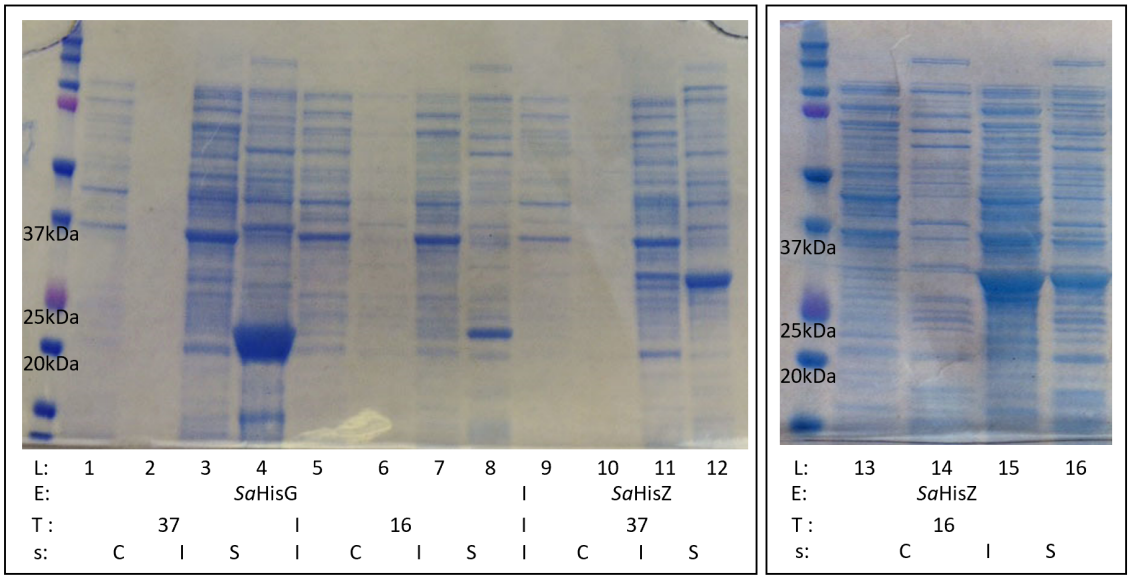


Figure 3.1: SDS-PAGE of *Sa*HisZ and *Sa*HisG$_S$ cell lysate from small scale expression tests of temperature and IPTG concentation. Odd lane numbers are soluble fractions and even lane numbers are insoluble fractions. *Sa*HisG$_S$ expression tests are in lanes 1–8 and *Sa*HisZ in lanes 9-16. Lanes 1–4 and 9–12 were expression tests carried out at 37 °C and lanes 5–8 and 13–16 at 16 °C. Lanes 1–12 used 1 mM IPTG and lanes 13–16 used 0.5 mM IPTG expression conditions. Lanes 1–2, 5–6, 9–10 and 13–14 were control samples pre-induction and lanes 3-4, 7-8, 11–12 and 15–16 were samples post-induction. Lane (L), enzyme (E), temperature (T), sample type (s), test sample (S), control sample (C). Molecular weight ladders are present in unmarked lanes with size label below corresponding protein band.

A large-scale purification and TEVP cleavage of *Sa*HisZ using 16 °C and 0.5 mM expression conditions was trialled. The initial nickel column was successful with a large amount of *Sa*HisZ accounted for on the gel (Figure 3.2, A). Unfortunately, the TEVP cleavage suffered from extensive crashing out of protein. Upon SDS-PAGE analysis it was determined that TEVP had crashed out and *Sa*HisZ remained in solution. Therefore, after filtering and re-cleaving, a second nickel column was run. Although some *Sa*HisZ was present in the wash, the majority was found in the imidazole gradient factions (Figure 3.2, B). This suggested that either, *Sa*HisZ had not fully cleaved (unlikely given SDS-PAGE showed a less intense band compared to uncleaved sample) or that *Sa*HisZ has some affinity to the column without the 6xHis-tag. A third nickel column using loading and wash buffer with 10 mM imidazole resulted in *Sa*HisZ eluting in the wash.

This confirms that the 6xHis-tag had now been successfully removed and that untagged *Sa*HisZ does have some affinity for the nickel column once the 6xHis-tag has been removed. However, there were still trace amounts of protein at lower molecular weight than *Sa*HisZ present in the protein solution. Gel filtration produced one clean peak but SDS-PAGE (Figure 3.3) of the resulting fractions highlighted lower molecular weight protein bands were still present. Mass spectrometry and a Mascot protein search of the NCBI database confirmed the extra bands to be *Sa*HisZ. This is likely to be an artefact of the gel rather than degradation products as the samples were re-run and no additional bands were present. Furthermore, the intact mass was 31710.4 Da by ESI-MS, matching the actual mass of cleaved *Sa*HisZ protein (31710.9 Da).

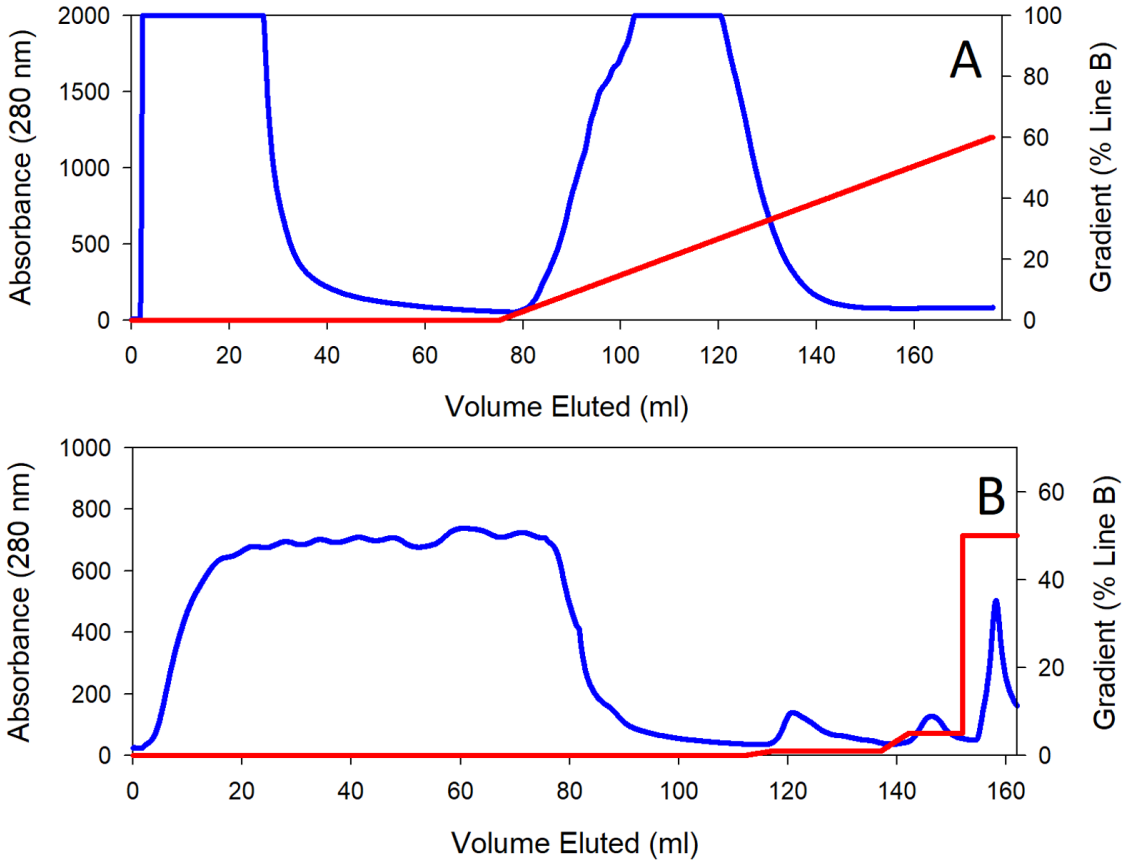


Figure 3.2: FPLC elution profile of *Sa*HisZ nickel column purification steps from A) pre-TEVP cleavage and B) post-TEVP cleavage.

Despite initial method setbacks, 72 mg of *Sa*HisZ was successfully purified. There was clear scope to improve the method for subsequent purifications but no further batches of *Sa*HisZ were ever required.

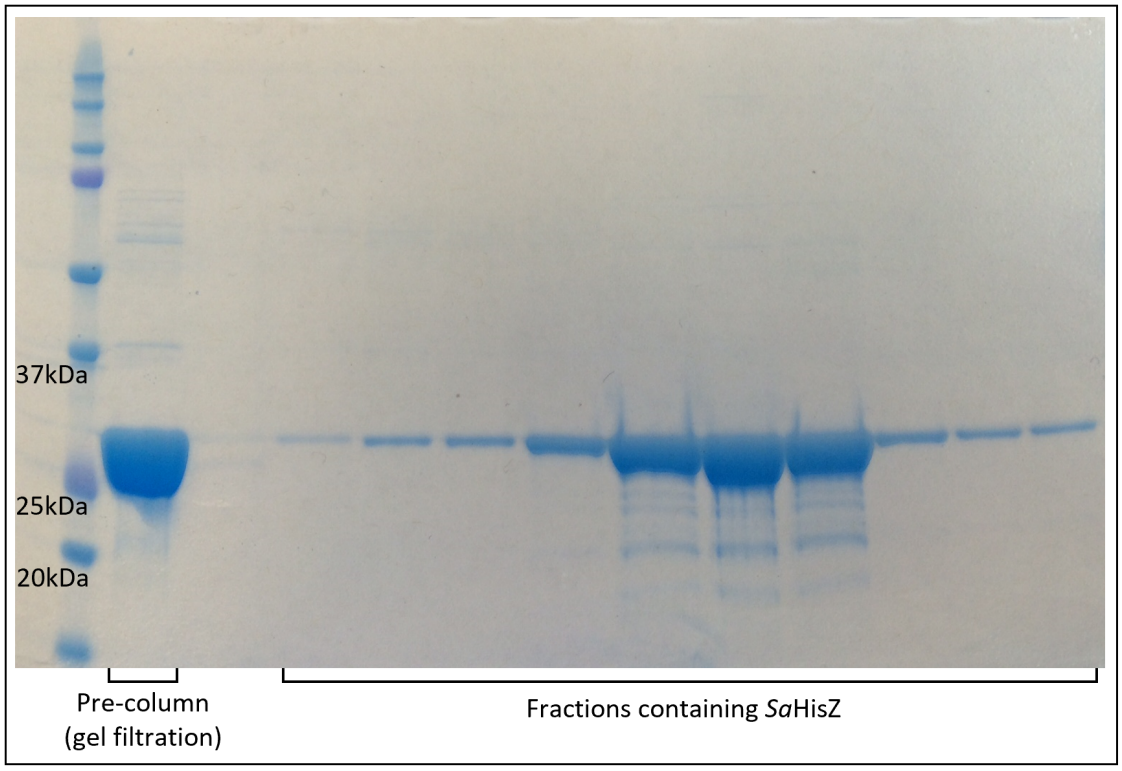


Figure 3.3: SDS-PAGE after final *Sa*HisZ purification step (untagged). Molecular weight ladder is present in left hand lane with size label immediately below corresponding protein band.

3.1.2 *Sa*HisG$_S$ Expression and Purification

All expression attempts from this section are summarised in Table 3.1.

Temperature and IPTG Concentration

Over-expression of *S. aureus* protein *Sa*HisG$_S$ (24403 Da with 6xHis-tag) was successful with 1 mM IPTG at both temperatures, 37 °C and 16 °C (Figure 3.1) but expressed in the insoluble fraction of the cell lysate. Upon successful soluble over-expression of *Sa*HisZ when IPTG concentration was lowered, expression trials with 0.5 mM and 0.1 mM IPTG were set-up for *Sa*HisG$_S$. Unfortunately, although there was more over-expression with 0.5 mM IPTG than 1 mM or 0.1 mM IPTG, the protein remained insoluble for all IPTG concentrations at 37 °C and 16°C.

Alternative Strains

Given that temperature and IPTG concentration changes had not resulted in soluble *Sa*HisG$_S$ in BL21(DE3), alternative strains were considered. Using C43 (DE3) cells, a strain specifically engineered for recombinant proteins which may be toxic to BL21(DE3) cells [112], over-expression of *Sa*HisG$_S$ was successful at 0.5 mM and 0.1 mM IPTG at 37 °C but all *Sa*HisG$_S$ resided in the insoluble fraction (Figure 3.4 A). No over-expression was found in C43 (DE3) at 16°C at either concentration of IPTG. Using Origami 2(DE3) cells, a strain with additional folding apparatus [113], *Sa*HisG$_S$ over-expressed at both temperatures and 0.1 mM and 0.5 mM of IPTG but all over-expressed protein was insoluble (Figure 3.4 C). The outcome was the same for BL21Gold(DE3) cells (not shown). These cells are superior strains of BL21(DE3) with two proteases capable of degrading recombinant protein removed [114]. Despite being insoluble, the best over-expression was achieved using BL21(DE3) and C43 (DE3). Another strain with advantages for leaky expression of toxic proteins was tested, BL21GoldPlysS(DE3)[**314**], but no expression was found, except at 37°C where the protein was insoluble (not shown).

It was hoped that SoluBL21 could help solubilise *Sa*HisG$_S$ since the strain is specifically optimised for this purpose [128]. Unfortunately, no over-expression of *Sa*HisG$_S$ was identified using SoluBL21 cells at 37 °C or 16 °C (Figure 3.4 B), even with an additional 24 hour growth period after induction (0.5 mM IPTG).

Table 3.1: Summary of *Sa*HisG-pJexpress411 expression results from different temperatures, IPTG concentrations, *E. coli* strains and media. BLGold refers to the strain BL21(DE3)Gold and BLGoldplysS refers to the strain BL21(DE3)Gold pLysS.

E.coli Strain	Media	Temp. (°C)	[IPTG] (mM)	Lysis Buffer	Expression	Soluble $SaHisG_S$
BL21(DE3)	LB	37	1	Bug Buster	✓	×
BL21(DE3)	LB	16	1	Bug Buster	✓	×
BL21(DE3)	LB	37	0.5	Bug Buster	✓	×
BL21(DE3)	LB	16	0.5	Bug Buster	✓	×
BL21(DE3)	LB	37	0.1	Bug Buster	✓	×
BL21(DE3)	LB	16	0.1	Bug Buster	✓	×
C43 (DE3)	LB	37	0.5	Bug Buster	✓	×
C43 (DE3)	LB	16	0.1	Bug Buster	×	×
C43 (DE3)	LB	37	0.5	Bug Buster	✓	×
C43 (DE3)	LB	16	0.1	Bug Buster	×	×
Origami 2	LB	37	0.5	Bug Buster	✓	×
Origami 2	LB	16	0.1	Bug Buster	✓	×
Origami 2	LB	37	0.5	Bug Buster	✓	×
Origami 2	LB	16	0.1	Bug Buster	✓	×
SoluBL21	LB	37	0.5	Bug Buster	×	×
SoluBL21	LB	16	0.5	Bug Buster	×	×
BLGold	LB	37	0.5	Bug Buster	✓	×
BLGold	LB	16	0.5	Bug Buster	✓	×
BLGold	LB	37	0.1	Bug Buster	✓	×
BLGold	LB	16	0.1	Bug Buster	✓	×
BLGoldplysS	LB	37	0.5	Bug Buster	✓	×
BLGoldplysS	LB	16	0.5	Bug Buster	×	×
C43 (DE3)	TB	37	0.5	Bug Buster	✓	×
C43 (DE3)	TB	16	0.5	Bug Buster	✓	×
C43 (DE3)	TPB	37	0.5	Bug Buster	✓	×
C43 (DE3)	TPB	16	0.5	Bug Buster	✓	×
C43 (DE3)	LB & Mg^{2+}	37	0.5	Bug Buster	✓	×
BL21(DE3)	Auto-ind.	16	N/A	Bug Buster	✓	×
BL21(DE3)	Auto-ind.	16	N/A	Emulsifier	✓	×

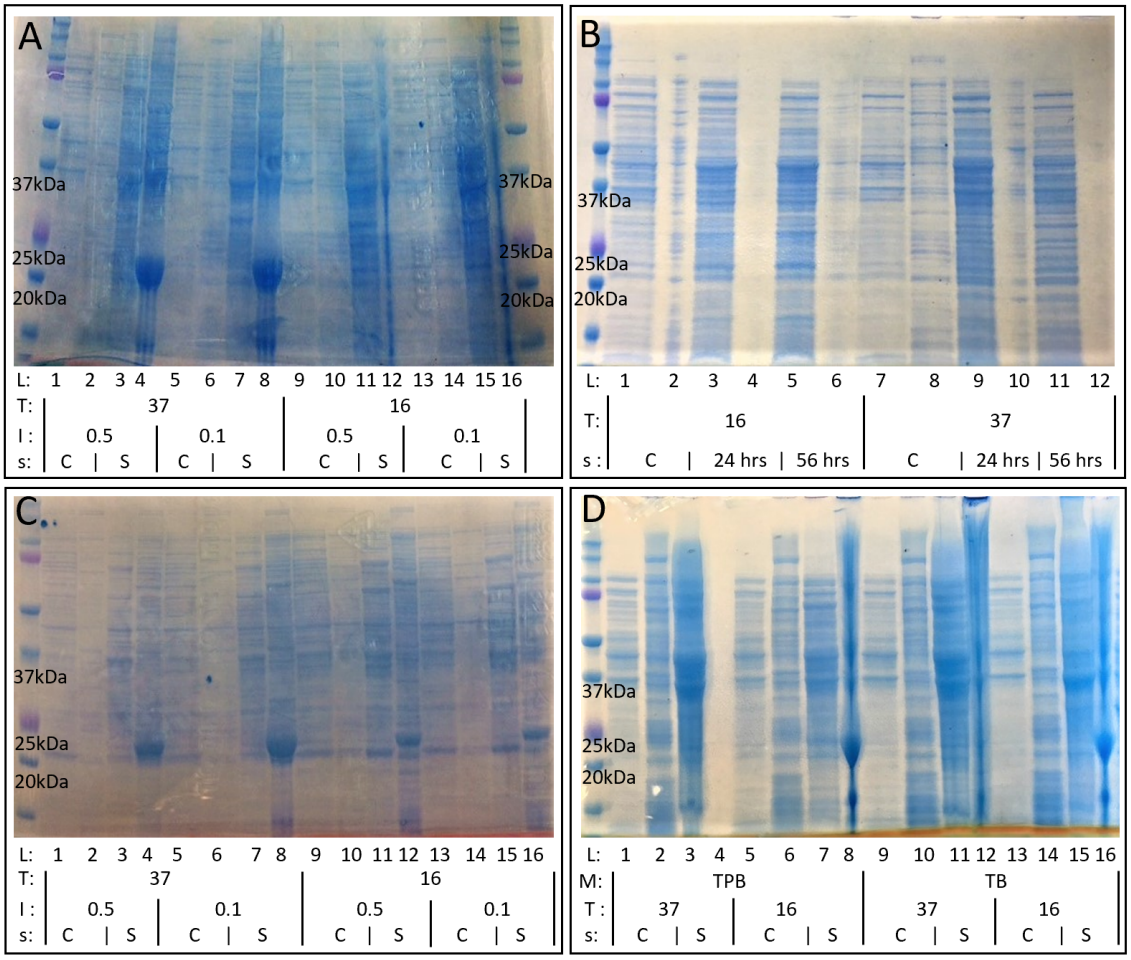


Figure 3.4: SDS-PAGE of *Sa*HisG$_S$ cell lysate from small scale expression tests of strains (A) C43(DE3) (B) SoluBL21 (C) Origami 2(DE3) and (D) TPB and TB media. Lanes (L) are numbered with temperature (T), IPTG concentration (I), media (M) and sample type (s), either control (C) or experimental sample (S) are indicated below each figure. Odd lane numbers are from soluble fractions and even lane numbers are from insoluble fractions. Molecular weight ladders are present in unmarked lanes with size label below corresponding protein band.

Alternative Media

Over-expression of *Sa*HisG$_S$ was only observed at 37°C for TPB and TB and only in insoluble form (Figure 3.4 B). Previous work by Yang et. al.[129] showed that adding magnesium ions to the media can increase solubility by 50 %. However, adding magnesium into LB at 37°C did not improve target protein solubility in this case.

Auto-induction

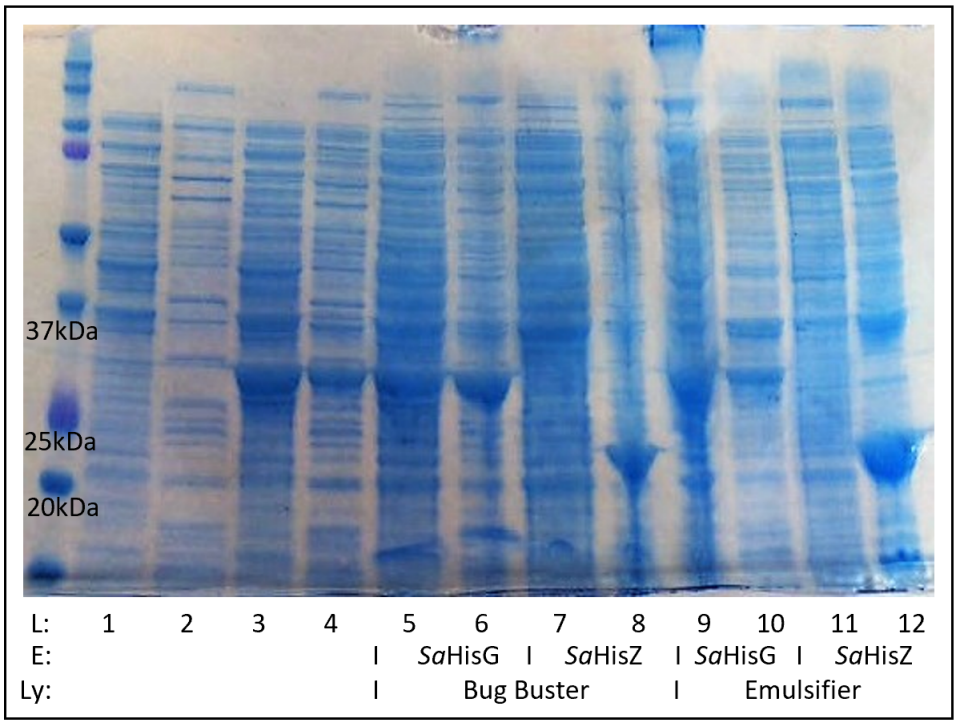


Figure 3.5: SDS-PAGE of *Sa*HisZ and *Sa*HisG$_S$ cell lysate from auto-induction expression tests. Odd lane numbers are soluble fractions and even lane numbers are insoluble fractions. *Sa*HisG$_S$ expression tests are in lanes 7-8 and 11-12 and *Sa*HisZ expression tests are in lanes 5-6 and 9-10. Lanes 5-8 used Bug Buster as the cell lysis method whereas lanes 9-12 used an emulsifier. Lane (L), enzyme (E), lysis type (Ly). Molecular weight ladders are present in unmarked lanes with size label below corresponding protein band. Note there are no control samples due to the nature of the method. Lanes 1-4 correspond to lanes 13-16 in Figure 3.1

.

An auto-induction method was tested with *Sa*HisG$_S$ and the resulting cell pellet lysed with either Bug Buster or by a high-pressure cell disruptor (emulsification). It was hoped the cell disruptor would aid solubilisation by breaking up any inclusion bodies that may have formed. *Sa*HisG$_S$ over-expressed using the auto-induction method but remained insoluble by both the Bug Buster and the emulsification methods (Figure 3.5). An auto-

induction method was also tested with *Sa*HisZ and although some soluble protein was obtained it was not an improvement upon the original expression conditions in Section 3.1.1.

3.1.3 *Sa*HisG$_S$-pJexpress411 Plasmid Engineering

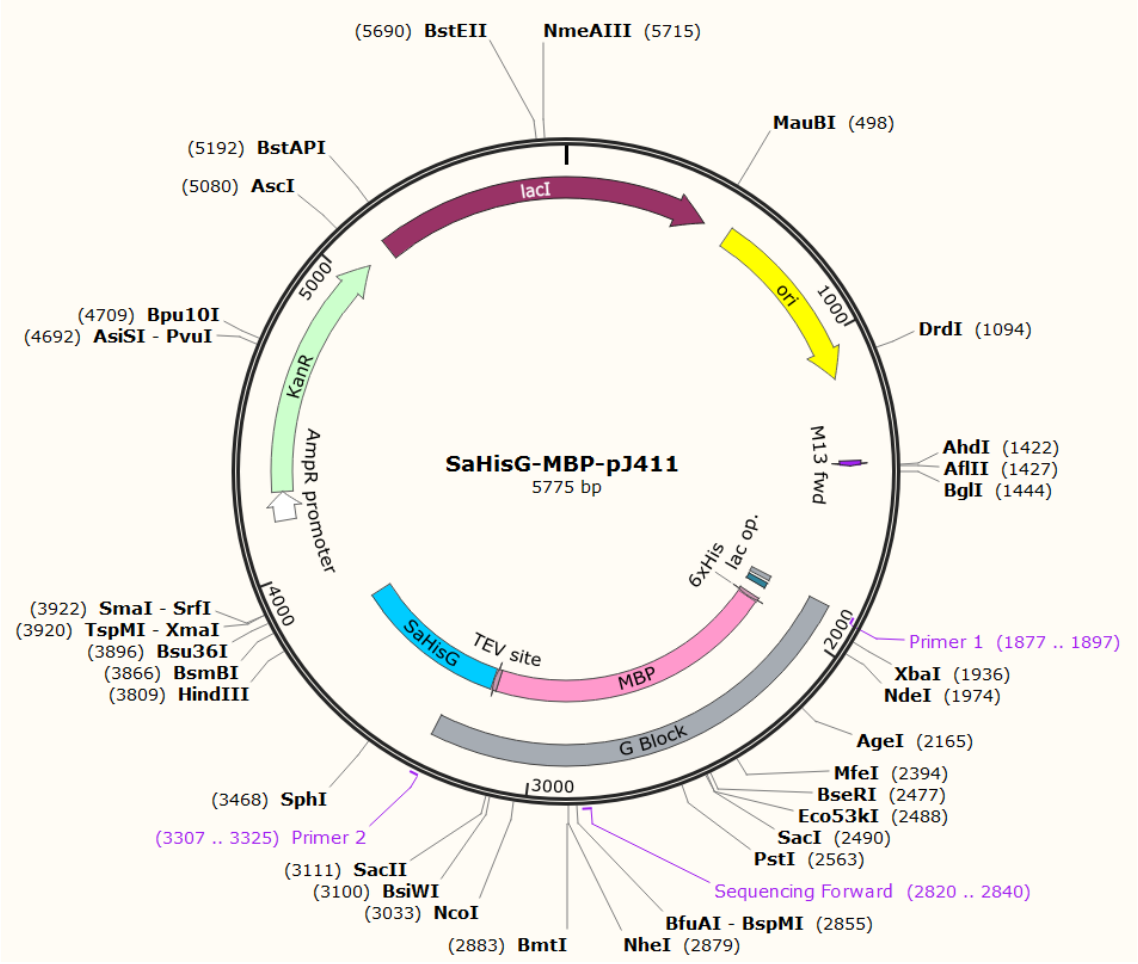


Figure 3.6: *Sa*HisG$_S$-MBP-pJ411 plasmid map. Restriction sites are shown in bold and primer regions used for cloning and sequencing are in purple text. KanR (green box) denotes the region containing a gene for kanamycin resistance and lacI (maroon box), the lacI gene. ORI (yellow box) is the origin of replication. *Sa*HisG gene is shown in the blue box and the MBP gene in the pink box. The G-Block that was inserted into the plasmid during cloning is shown in the grey box. Image was obtained from SnapGene software.

As a technique to improve *Sa*HisG$_S$ solubility, maltose binding protein (MBP), a particularly soluble protein [130], was inserted into the *Sa*HisG$_S$-pJ411 plasmid between the 6xHis-tag and TEV cleavage site (see Figure 3.6) using the Gibson Assembly method. Initially, PCR was used to amplify linearized fragments of the original *Sa*HisG$_S$-pJ411

plasmid from Primer 1 anti-clockwise to Primer 2 (can be visualised using Figure 3.6). After the PCR, *Dpn*1 (which only cleaves when its recognition site is methylated) was used to remove any unamplified parent plasmid. Gibson Assembly was used to insert a G-Block fragment (IDT) containing the MBP gene into the linearized plasmid. Gibson assembly uses a T5 exonuclease to chew back the PCR fragment and G-block ends. Complementary fragments anneal and a DNA polymerase completes the strands. Finally, a DNA ligase seals the nicks and the plasmid cloning is complete. It was hoped that MBP would increase the overall solubility of the protein and bring *Sa*HisG$_S$ into the soluble fraction. The MBP would then be cleaved off along with the 6xHis-tag during dialysis with TEVP. There is a risk that *Sa*HisG$_S$ would be dragged into the soluble fraction by MBP and remain insoluble upon TEVP cleavage, but others [131, 132] have shown success with this strategy.

PCR to Linearize *Sa*HisG$_S$-pJ411 Fragment

Using primers designed to amplify the region of the plasmid which didn't contain the G-Block fragment, PCR was used to linearize the original *Sa*HisG$_S$-pJ411 plasmid.

Of the eight conditions tested, only those containing no DMSO were successful with bands at ~4000 bp being present at annealing temperatures of 53.2 °C to 55 °C (Figure 3.7, red box). This corresponds to the size of the linear plasmid fragment (4039 bp) expected but the bands were very weak. There was also low bp contaminants present under these conditions however it is unlikely that they carried the full sequence for antibiotic resistance and so are unlikely to pose problems during further processing (as any colonies with those DNA will not grow once transformed). The sample from lane 5 (annealing temperature of 53.2 °C and containing DMSO) was selected for the Gibson Assembly since it had the most intense 4000 bp band on the agarose gel.

Gibson Assembly of Linear *Sa*HisG$_S$-pJ411 with MBP G-Block

Of the two G-Block concentrations tested (5x and 10x), the 10x G-Block condition produced the most viable colonies after transformation (22 colonies compared to 5 colonies for 5x). This is plausible since the 10x condition has more G-Block fragments available to insert into the linear plasmid fragments. Overall the small number of viable colonies is likely a direct result of only a weak band having been present after the PCR. This suggests that there was only a small amount of linear plasmid fragment available for the assembly. Additionally, dead end assembly may have occurred with the contaminant fragments that were also present after the PCR.

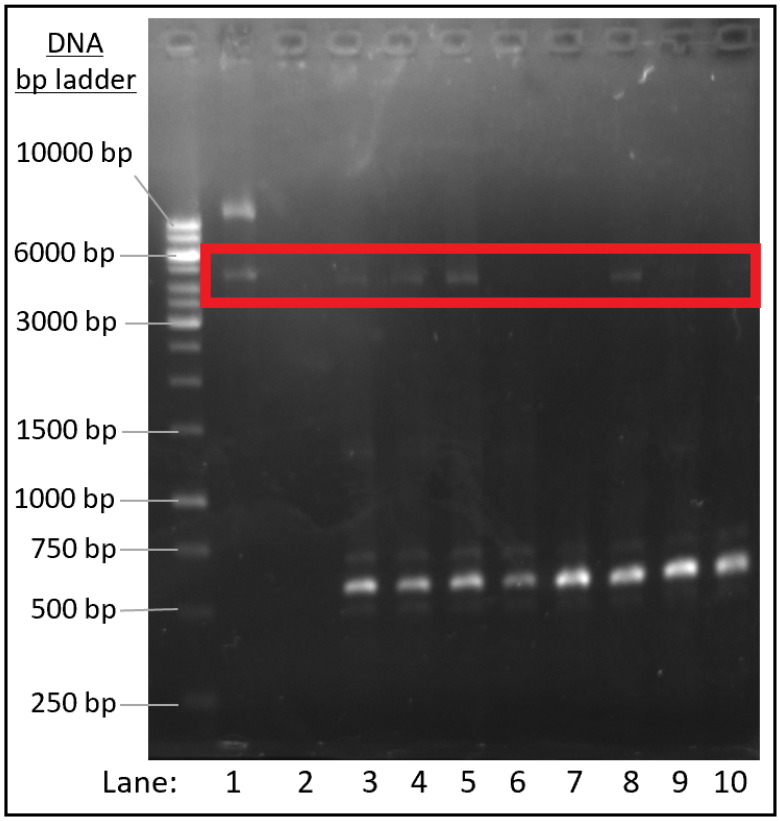


Figure 3.7: Agarose gel of PCR to linearize SaHisG-pJ411. Lanes 7–10 contained DMSO, lanes 3–5 contained no DMSO. Annealing temperature was 52 °C for lanes 10 and 6, 53.2 °C for lanes 5 and 9, 54 °C for 8 and 4, and 55 °C for lanes 2, 3 and 7. Lane 2 was a control lacking DNA polymerase and lane 1 was non-linearized SaHisG-pJ411. The un-numbered lane is a Thermo Fisher Gene Ruler 1 KB DNA ladder with length corresponding to those annotated.

Plasmid Purification of Gibson Assembly DNA

If the Gibson Assembly was successful, the purified plasmids would contain a *HindIII* site and an *NdeI* site which after digestion with both of these enzymes, would yield fragments of 1834 bp and 3941 bp long.

After digestion with *Hind*III and *Nde*I both plasmid M and plasmid X produced fragments ~4000 bp and ~1800 bp long (Figure 3.8). This is good initial evidence that the Gibson Assembly was successful and the MBP gene has been inserted into the *Sa*HisG$_S$-pJ411 plasmid (exact size 5775 bp). The undigested plasmid band for both plasmids, however, were at ~10000 on the gel. Given that there are no other bands on the gel to account for the rest of the base pairs after digestion, and that there is streaking on the band it is likely that the band ran higher due to concentrated circular DNA.

Furthermore, the original *Sa*HisG$_S$-pJ411 plasmid was 4599 bp and the fragments

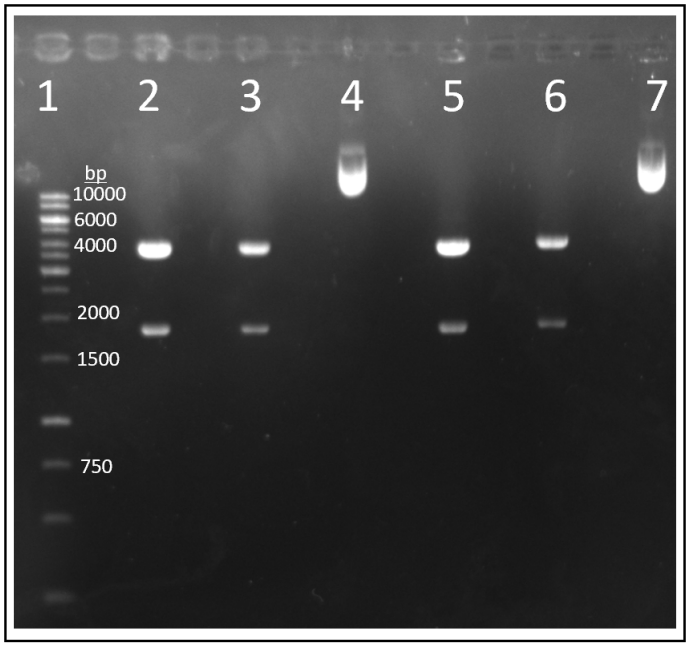


Figure 3.8: Agarose gel of Gibson Assembly digestion of plasmid with HindIII and NdeI. Lane 1 is a base pair ladder as labelled, 2 is plasmid X (high concentration digest), 3 is plasmid X (low concentration digest), 4 is plasmid X with no digestion, 5 is plasmid M (high conconcentraion digest), 6 is plasmid M (low concentration digest), 7 is plasmid M with no digestion.

resulting from *Hind*III and *Nde*I digestion would be 3940 bp and 659 bp. There is no evidence of the small 659 bp fragment on the gel and so the fragment bands cannot merely be from digestion of any reformed or residual parent plasmid. This should not be possible either due to the *Dpn*I digestion after plasmid linearization which should remove any parent DNA left after PCR.

3.1.4 Expression Tests of *Sa*HisG$_S$-pJ411-MBP

All expression test arbitrarily began with *Sa*HisG$_S$-MBP-pJ411 Construct X (*Sa*HisG$_S$-MBP(X) while awaiting sequencing results of both construct X and construct M. The expected size of *Sa*HisG$_S$-MBP(X) is 67302 Da.

Temperature, IPTG concentration and Induction Time

Despite the soluble MBP protein having been inserted into the *Sa*HisG$_S$-pJ411 plasmid, no soluble expression was found at 15, 20 or 37 °C in BL21(DE3) (Figure 3.9). It appears the MBP protein is still not capable of overcoming the insolubility of *Sa*HisG$_S$

in BL21(DE3), even if the post-induction growth time was 2 days or 0.1 mM IPTG was used.

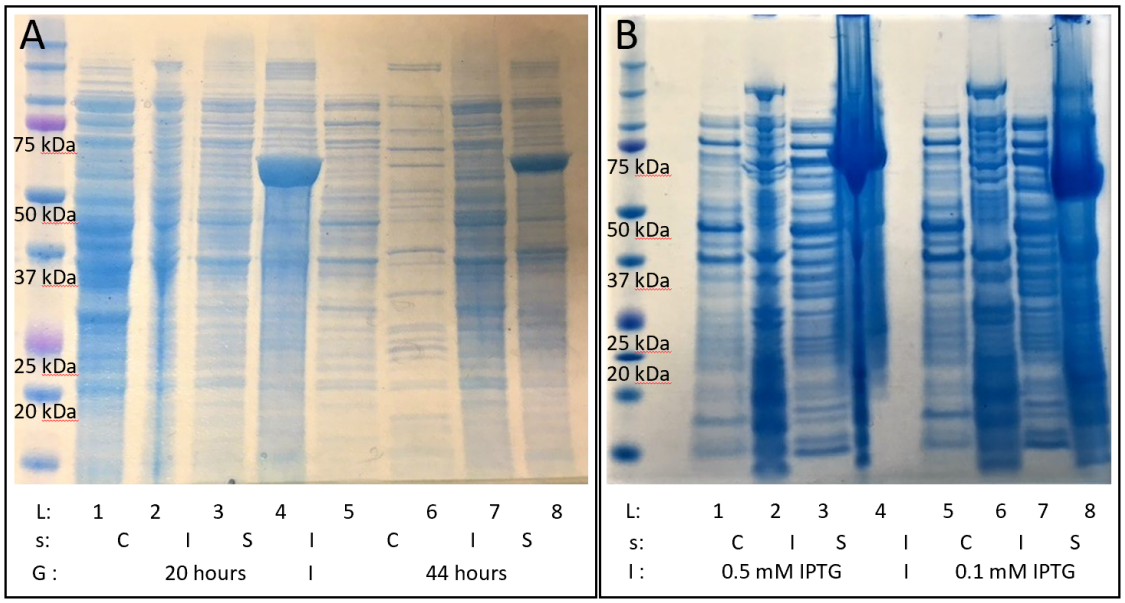


Figure 3.9: SDS-PAGE of SaHisG$_S$-MBP(X) cell lysate from small scale expression tests of (A) temperature and post-induction growth time at 37 °C and (B) IPTG concentration at 20 °C . Molecular weight ladders are present in unmarked lanes with size label below corresponding protein band. Odd lane numbers are soluble fractions and even lane numbers are insoluble fractions. Lanes 1–2 and 5–6 were control samples and lanes 3–4 and 7–8 were test samples. In (A), lanes 1–4 had post induction growth times of 20 hours and lanes 5–8 had 44 hours. In (B) lanes 1–4 used 0.5 mM IPTG and lanes 5–8 used 0.1 mM IPTG. Lane (L), sample type (s), control sample (C), test sample (S), IPTG concentration (I), post-induction growth time (G).

E. coli Strains

No over-expression (soluble or insoluble) of *Sa*HisG$_S$-MBP(X) was found using C43(DE3) or SoluBL21 cells (Figure 3.10), despite varying temperature and post-induction growth time (not shown). Over-expression of *Sa*HisG$_S$-MBP(X) did occur in Origami 2 cells at 37°C but the target protein was in the insoluble fraction. This is a good indication that the plasmid has been successfully modified as all components were present for an over-expression, however, in this case, inserting MBP into the plasmid has not benefitted *Sa*HisG solubility.

Given that even if *Sa*HisG$_S$-MBP(X) was to move across to the soluble fraction under different conditions, there is nothing to ensure that *Sa*HisG$_S$ would not crash out after cleavage with TEVP to remove the 6xHis-tag and MBP from the target. Due to this, no

further *Sa*HisG$_S$-MBP conditions were tested.

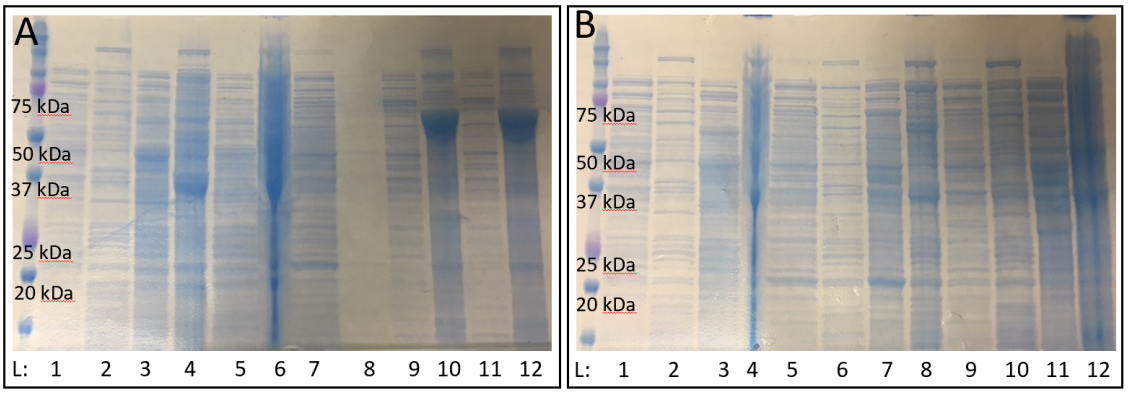


Figure 3.10: SDS-PAGE of *Sa*HisG$_S$-MBP(X) cell lysate from small scale expression tests of SoluBL1, Origami 2 and C34(DE3) at (A) 37 °C and (B) 16 °C. Molecular weight ladders are present in unmarked lanes with size label below corresponding protein band. Odd lane (L) numbers are soluble fractions and even lane numbers are insoluble fractions. Lanes A1–2, A7–8, B1–2, B5–6 and B9–10 were control samples and lanes A3–6, A9–12, B3–4, B7–8 and B11–12 were test samples. Lanes A1–6 and B1–4 used SoluBL21 cells, A7–12 and B5–8 used Origami2(DE3) cells and B9–12 used C43(DE3) cells.

Sequencing of *Sa*HisG$_S$-MBP -pJ411

The *Sa*HisG$_S$-MBP-pJ411 (constructs X and M) were sent for commercial sequence analysis (Eurofins Genomics) to ensure that that modified plasmid contained the full MBP sequence. The sequencing analysis confirmed a 736 base pair stretch correlating to the desired *Sa*HisG$_S$-MBP-pJ411 sequence in construct M. Furthermore, this span confirmed the successful joining of the MBP G-block and the original plasmid *Sa*HisG$_S$ gene.

Unfortunately, with the *Sa*HisG$_S$-MBP(X)-pJ411 plasmid, only a 174 base pair stretch was confirmed and did not include the G-block and *Sa*HisG$_S$ join. Since the X construct was used for all of the expression tests it is possible that they are not a true representation of *Sa*HisG$_S$-MBP-pJ411. However, over-expression was observed in Origami 2 cells which does indicate that the plasmid is fully inducible. Repeat expression tests using Construct M could probe the Construct X results but this was deemed unlikely to provide fruitful results.

3.1.5 Large Scale *Sa*HisG$_S$-pJ411 Protein Purification Solubility Trials

It was hoped that *Sa*HisG$_S$ (if present at levels lower than that visible by SDS-PAGE from 1 ml of culture) would concentrate on the column and become visible when eluted under the gradient conditions of the column. However, this was not the case for *Sa*HisG$_S$-pJ411 in C43(DE3), both of which were unsuccessful under the conditions tested. Potential bands, at a similar height on the gel to where *Sa*HisG$_S$ was expected (~24kDa,) were sent for mass spectrometry identification but were subsequently identified as general *E. coli* proteins (not shown).

3.1.6 *Sa*HisG$_S$-pJ411 Tris/Bis Tris Lysis Buffer Solubility Trials

Alternative lysis buffers were tested to determine whether poor solubility was due to *Sa*HisG$_S$ instability in Bug Buster or Buffer A1. Bis-Tris and Tris buffers were trialled with different additives including amino acids, reducing agents and detergents. A selection of the results are shown in Figure 3.11.

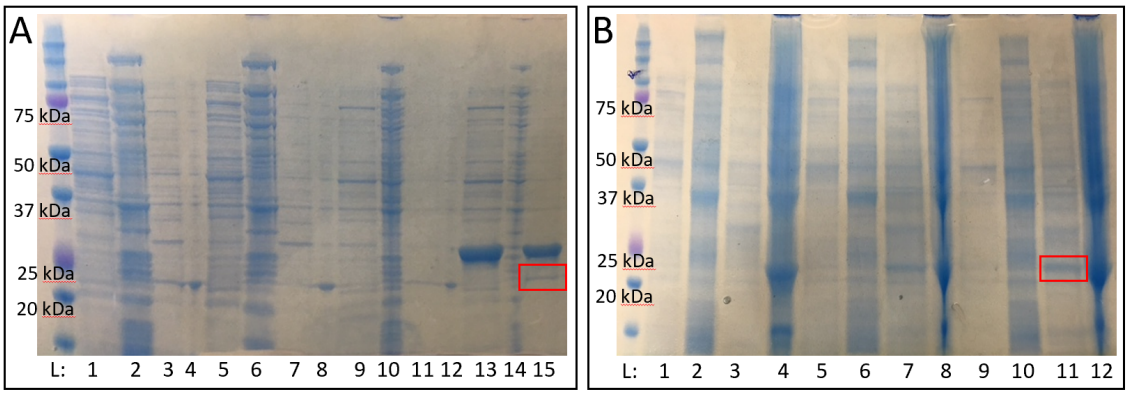


Figure 3.11: SDS-PAGE gel from (A) Tris and (B) Bis-Tris buffer screen for soluble *Sa*HisG$_S$. Molecular weight ladders are present in unmarked lanes with size label below corresponding protein band. Odd lane (L) numbers are soluble fractions and even lane numbers are insoluble fractions. Lanes 1–2, 5–6, 9–10, 13–14 were control samples and lanes 3–4, 7–8, 11—12 and 15 were test samples. Lanes 1–4 contained Igepal CA360, 5–8 contained Triton X100, 9–12 contained β-mercaptoethanol and lanes 13–15 contained *Sa*HisZ. Red boxes indicate soluble *Sa*HisG$_S$ bands. Note that some wells ran poorly and were repeated (not shown).

Overall the alternative buffers performed less well than Bug Buster at protein extraction (Figure 3.1), but this is to be expected since Bug Buster is specifically formulated for this purpose. From the Tris buffers screened only the buffer containing purified *Sa*HisZ showed a band by SDS-PAGE that may be attributed to soluble *Sa*HisG$_S$ (Figure 3.11 A, red square). It was hoped that *Sa*HisZ would pull *Sa*HisG$_S$ into the soluble fraction

by forming holoenzyme complex. Well overflow contamination from the neighbouring insoluble sample was ruled out by repeating with a blank lane between each sample (not shown). A band at ~24 kDa was still present in the soluble fraction. This was confirmed to be *Sa*HisG$_S$ by MALDI mass spectrometry. Since *in vivo* *Sa*HisG$_S$ and *Sa*HisZ associate with each other to form the ATPPRT complex it is likely that upon cell lysis the *Sa*HisZ in the buffer helps to stabilise soluble *Sa*HisG$_S$. This effect is also seen when *Sa*HisZ is added to a Bis Tris buffer. See Section 3.1.8 for further *Sa*HisZ conditions tested.

Of the Bis Tris buffers tested, two more potential lysis buffers leading to soluble *Sa*HisG$_S$ were noted (Figure 3.11 B). The addition of Triton X100 led to a protein band at ~24 kDa that was not present in the pre-induction control samples. Detergent is commonly used to aid protein solubility however they are not compatible with many analytical techniques, including mass spectrometry. Due to this, a buffer containing detergent is not desired. However, the addition of the reducing agent β-mercaptoethanol (BMCE) to Bis-Tris also resulted in a protein band at ~24kDa that was not present in the pre-induction controls (Figure 3.11 B, red square). This was repeated to confirm no cross contamination had occurred (not shown). It is likely that reducing agents such as BMCE are capable of stabilising *Sa*HisG$_S$ by reducing disulphide bonds that may form within the protein. Unfortunately, Bis-Tris is not compatible with the HisTrap FF column used for large scale purification and Tris (which is compatible) did not produce soluble *Sa*HisG$_S$.

3.1.7 *Sa*HisG$_S$-pJ411 Reducing Agent Buffer Solubility Trials

Reducing agents BMCE and DTT were tested with buffers compatible with a HisTrapFF nickel column. Overall buffers containing DTT produced a larger amount of soluble *Sa*HisG$_S$ than BMCE (Figure 3.12). DTT is a more desirable reagent since it is a solid at room temperature and less volatile but when added to Buffer A1 (50 mM HEPES, 500 mM NaCl, 10 mM Imidazole, pH 8) no soluble *Sa*HisG$_S$ was obtained. The protein could be salting out due to the high salt concentration or *Sa*HisG$_S$ may be unstable in imidazole. This was investigated using small scale expression tests to rule out each of the buffer components (not shown). Only the buffer containing 50 mM HEPES and 4 mM DTT resulted in soluble *Sa*HisG$_S$ which was confirmed by ESI mass spectrometry.

A large-scale purification using only HEPES, DTT and imidazole (no salt) in the buffer was tested but *Sa*HisG$_S$ was not collected from any of the column fraction elutions, nor was it present by eye in the pre-column sample on the resulting SDS-PAGE gel (not shown). It appears that lack of soluble *Sa*HisG$_S$ could be due to both salting out and imidazole instability. Histidine was tested as an imidazole replacement but with no

positive result (not shown).

3.1.8 *Sa*HisG$_S$-pJ411 Lysis Buffer Trial with *Sa*HisZ in Buffer

Tris and Buffer A1 components were tested with DTT in the presence of *Sa*HisZ in order to determine whether *Sa*HisZ could stabilise soluble *Sa*HisG$_S$ by forming the holoenzyme when imidazole was present in the buffer (Figure 3.13). However, it did not appear that *Sa*HisZ could dramatically alter the soluble stability of *Sa*HisG$_S$ in any buffers containing imidazole, regardless of DTTs presence.

3.1.9 Large Scale *Sa*HisG$_S$ Protein Purification Trials with DTT in Buffer

Being unable to use imidazole or histidine in the buffer for protein purification is not ideal but can be worked around. For example, imidazole could be removed from Buffer A but still be present in elution Buffer B. This would result in more unspecific proteins binding to the column initially but would leave *Sa*HisG$_S$ in a compatible buffer for the entirety of the preparation and binding to the column. Alternatively, pH could be used to elute *Sa*HisG$_S$ from the column. The first of these conditions was tested using large scale protein purification, however *Sa*HisG$_S$ was not visible in any of the elution fractions. Additionally, there was no obvious band present in the pre-column sample that was subsequently missing in the flow though to suggest *Sa*HisG$_S$ had been loaded and bound.

To investigate the whether *Sa*HisG$_S$ had been loaded onto the column at all, a 20–25 kDa section of SDS-PAGE gel of pre-column sample was analysed by mass spectrometry. The gel band was found to contain no *Sa*HisG$_S$. The only difference between the expression test where soluble *Sa*HisG$_S$ was present and the purification procedure (where it was not) was the scale up sample preparation. The expression test uses agitation in the lysis buffer to break open the cells to release *Sa*HisG$_S$ whereas the protein purification method utilises a cell emulsifier. The emulsifier should be more efficient and break up inclusion bodies but, in this case, it appears it could be be de-stabilising the soluble protein. It is possible that the protein is denatured due to the high pressure forced upon it in the emulsifier chamber.

3.1.10 Co-purification of *Sa*HisG$_S$ with *Sa*HisZ

A large-scale co-purification of *Sa*HisG$_S$ with *Sa*HisZ was trialled to determine whether formation of the holoenzyme from the beginning of cell paste processing could stabilise

*Sa*HisG$_S$ enough to withstand lysis by emulsifier and remain soluble. By using a large scale purification it was hoped that enough *Sa*HisG$_S$ would be concentrated on the column for a clear band to be present on the SDS-PAGE gel after elution. Conditions using 150 mM NaCl or 100 mM KCl were tested. Both purified *Sa*HisZ well, but neither gave measurable amounts of *Sa*HisG$_S$. Unfortunately it does not seem that *Sa*HisZ was able to stabilise *Sa*HisG$_S$ enough to withstand emulsification.

Emulsification is a relatively harsh cell lysis technique, particularly when compared to the small expression tests which used agitation with lysis buffer to lyse the cells. Unfortunately, it would be highly inefficient to attempt cell lysis by agitation for litres of cells (opposed to 1 ml). Therefore, lysis using a much milder lysis technique, shaking with glass beads was trialled instead. Unfortunately, no *Sa*HisG$_S$was present in the resulting elutions. Gel bands corresponding to 23-25 kDa from each elution step (flow through and gradient) were tested by mass spectrometry for traces of *Sa*HisG$_S$ but no related protein was found.

3.1.11 Concluding Remarks Pertaining to the Purification of *Sa*ATPPRT

Despite extensive attempts to solubilise *Sa*HisG$_S$ separately from *Sa*HisZ, none have been successful. Purifying both subunits separately was desirable for precise calculation of ATPPRT complex concentration (Equation 2.4) during kinetic studies, for ease of radioactive labelling in later experiments and for alternative projects planned within the group. Since separate purification was not successful, preparation to co-express *Sa*HisG$_S$ and *Sa*HisZ together using a pETduet vector (Novagen) had begun. It was hoped that by expressing both together, the *Pa*ATPPRT complex would form and *Sa*HisG$_S$ would be stabilised in soluble form. However, concurrently, Ghssein et al. [133] published the Biosynthesis of a Broad-spectrum Nicotianamine-like Metallophore in *Staphylococcus aureus*. Here, the authors presented identification of a new transport system, previously unidentified in *S. aureus*, containing a full ABC transporter suite capable of importing amino acids, including histidine, into the cell. Consequently, if the cell was deficient in histidine, for example from a corrupted histidine pathway, *S. aureus* would be able to scavenge histidine from the host. It is unknown whether *S. aureus* would be able to scavenge enough histidine from the host alone to overcome a compromised histidine biosynthesis pathway but, given there are currently no *in vivo* knock out studies proving otherwise, the enzymes in the *S.aureus* histidine pathway may not be as suitable a drug target as originally thought. This, coupled with extensive solubility problems, led to the cessation of the *Sa*ATPPRT project.

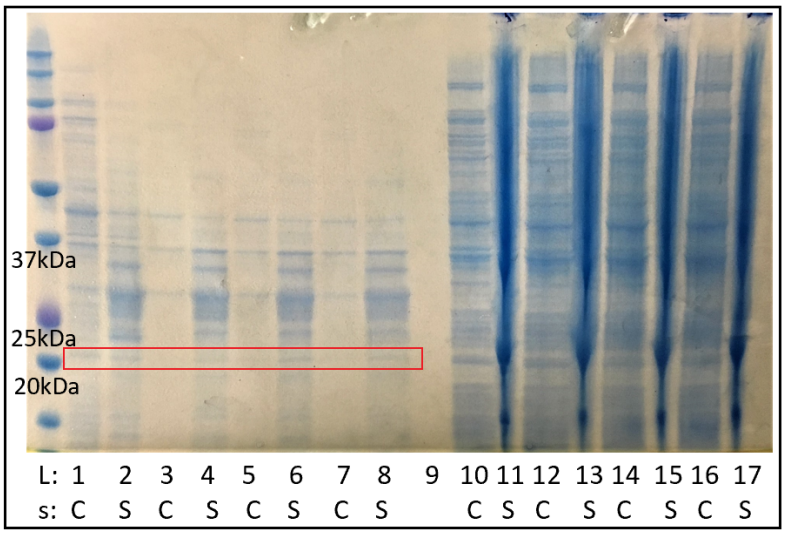


Figure 3.12: SDS-PAGE gel of reducing agent lysis buffer screen for soluble *Sa*HisG$_S$. Molecular weight ladders are present in unmarked lanes with size label below corresponding protein band. Sample type (s) are annotated as control samples (C) and test samples (S). Lanes (L) 1-8 were soluble samples and lanes 10-17 were insoluble samples. The lysis buffer used in Lanes 1-2 and 10-11 were HEPES and BMCE, lanes 3-4 and 12-13 were Buffer A1 and BMCE, lanes 5-6 and 14-15 were Buffer A1 and DTT, and lanes 7-8 and 16-17 were Buffer A1 alone. Lane 9 was intentionally left empty.

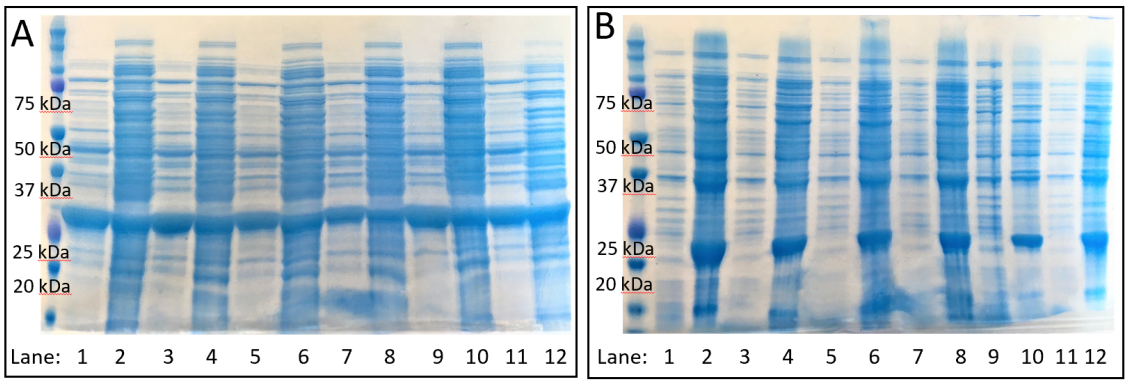


Figure 3.13: SDS-PAGE gel of a histidine and imidazole buffer screen for soluble *Sa*HisG$_S$. Molecular weight ladders are present in unmarked lanes with size label below corresponding protein band. Odd numbered lanes were control samples and even numbered lanes were test samples. The lysis buffer used in Lanes 1-2 were Tris, NaCl and *Sa*HisZ, lanes 3–4 were Tris, NaCl, *Sa*HisZ and imidazole, lanes 5–6 were HEPES, NaCl and *Sa*HisZ, lanes 7–8 were HEPES, NaCl, *Sa*HisZ and imidazole, lanes 9–10 were HEPES, NaCl, *Sa*HisZ and DTT, lanes 11–12 were HEPES, NaCl, *Sa*HisZ and histidine. (A) shows soluble fractions and (B) shows insoluble fractions.

3.2 Expression and Purification of *Pa*ATPPRT

3.2.1 Purification of *Pa*HisG$_S$ and *Pa*HisZ

Purification of *Pa*HisG$_S$ (25213 Da) and *Pa*HisZ (43085 Da) were both carried out as previously described [20] utilising two nickel column purification steps, one before and after TEVP cleavage of the 6xHis-tag (Figure 3.14). Initially, purification was as expected with purity >95% (Figure 3.15) and yields in the region of 30–40 mg L^{-1} for both *Pa*HisG$_S$ and *Pa*HisZ.

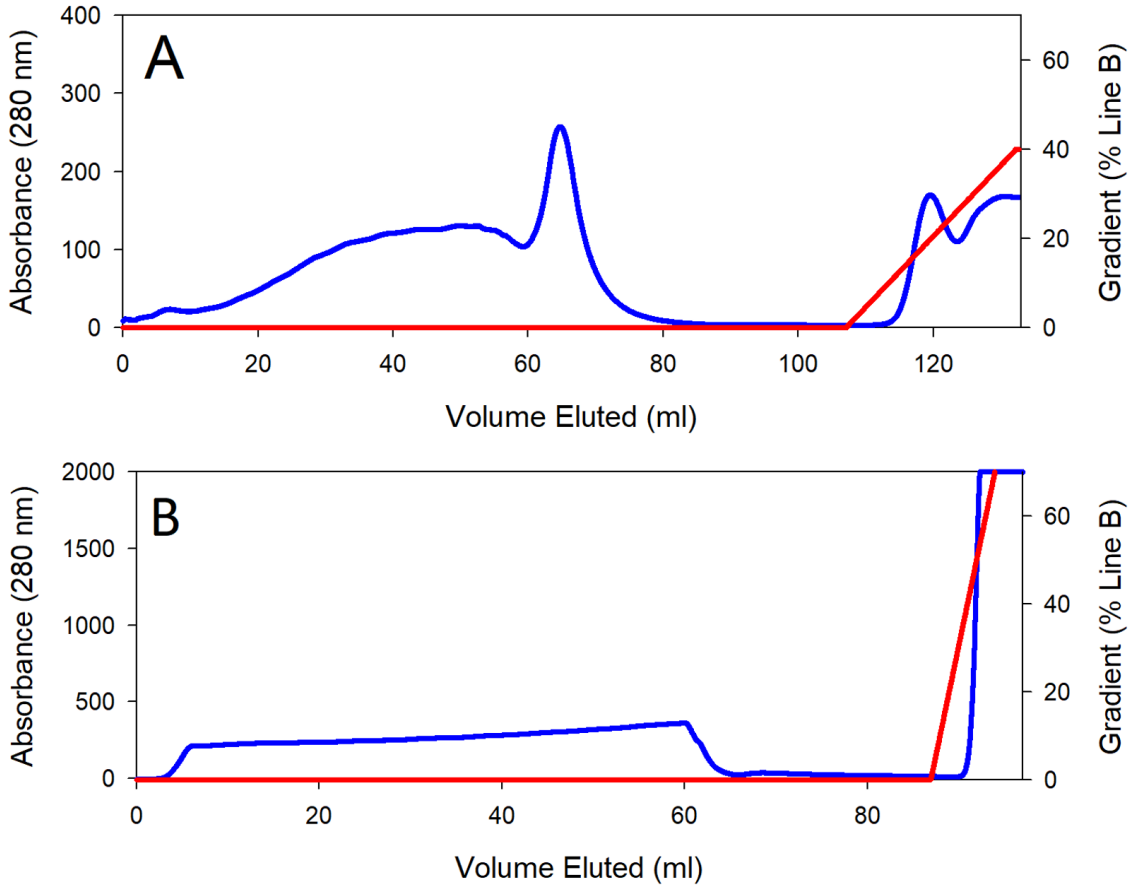


Figure 3.14: FPLC elution profile of *Pa*HisZ second nickel column purification step after TEVP cleavage from A) *Pa*HisG and B) *Pa*HisZ.

3.2.2 Method Development of Larger Scale *Pa*HisZ Purification

A larger scale purification was trialled, utilising a 20 ml nickel HisTrap column and 5 L of cells compared to the standard 5 ml column and 2 L of cells. It was hoped that a greater amount of protein could be obtained in a similar time period since a large quantity of protein is required for 'pre-steady state' kinetics. Unfortunately, using the larger column

did not improve purification attempts. *Pa*HisZ with >95% purity was obtained but with a yield of 90 mg total (18 mg L^{-1}) where as the standard purification protocol (using a column 5x smaller), obtains a greater amount of protein per litre of cells (30-40 mg L^{-1}) in substantially less time. The purification suffered from protein crashing out multiple times, a third chromatographic step and an extra 4 weeks purification time. Method development for a larger scale process was abandoned at this point in favour of smaller batch purification.

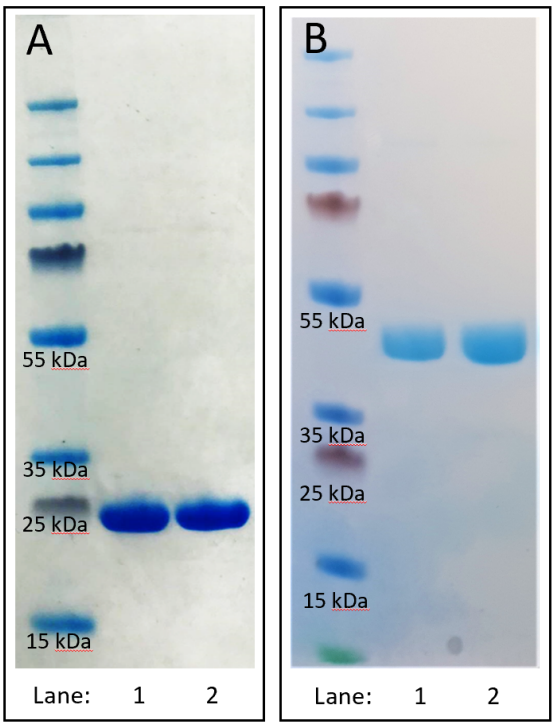


Figure 3.15: SDS-PAGE of final purity (A) *Pa*HisG$_S$ and (B) *Pa*HisZ purification. The first lane in both gels correspond to molecular weight marker, size as annotated directly underneath. Lanes labelled 1 and 2 are duplicates.

3.2.3 Troubleshooting *Pa*ATPPRT Purification Issues

Unfortunately, for a sustained period of time, purification was resulting in enzymes with considerably reduced activity (~4-fold drop). Initially it was thought that the pH of the assay buffer may have deviated from the required, pH 8.5. Figure 3.16 shows an example of the change in rate based on pH of the assay buffer. Figure 3.16 also shows the differences in rate observed for an original, fully active batch of *Pa*HisG$_S$ (blue) compared to one of the poor batches (orange). One of the differences between the good and poor quality batches was concentration. Although previously, batches of protein had been concentrated successfully to > 500 μM, the large difference in concentration between working batch 1 (19 μM) and compromised batch 2 (420 μM) led to troubleshooting of the

concentration process. Additionally, the concentrating process became under scrutiny due to observation that the poor batch accumulated many 'microbubbles' upon the surface when concentrating (a characteristic of denatured protein). New ultrafiltration membranes were used with the ultrafiltration concentrator, but this did not improve activity, nor did switching to using a centrifugal concentrator, or leaving batches very dilute.

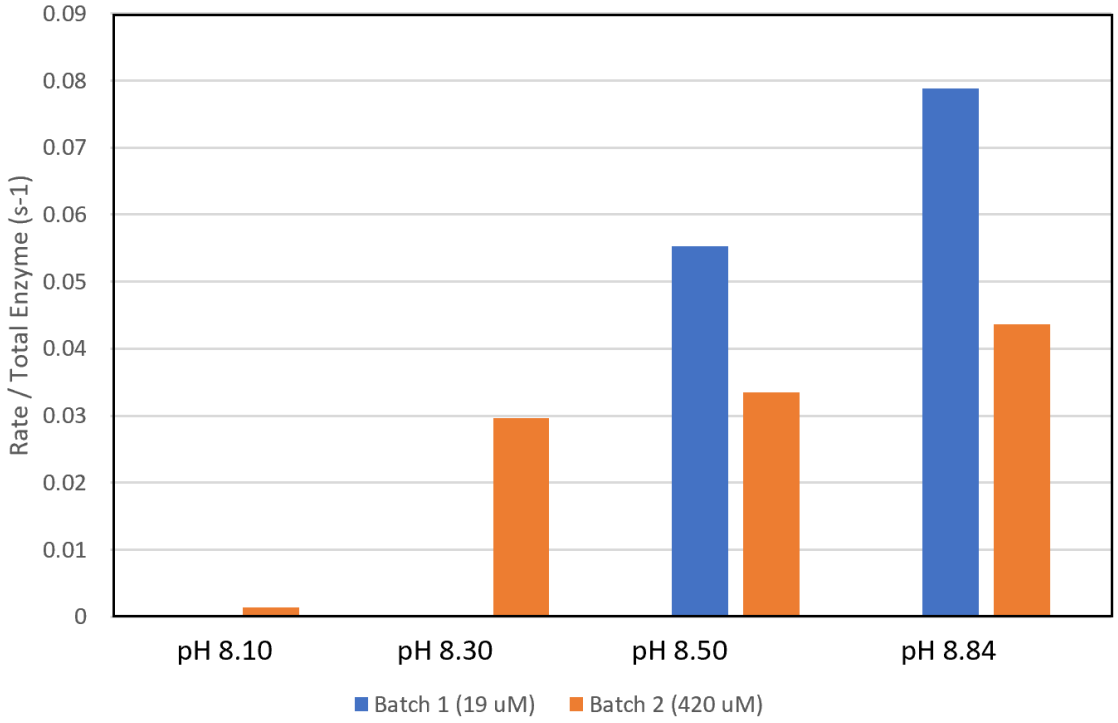


Figure 3.16: pH dependence on the rate of *Pa*HisG$_S$ activity corrected for total enzyme used. Batch 1 is one of the original, fully active batches (blue) and Batch 2 is one of the poor rate batches of *Pa*HisG$_S$ (orange).

Differential Scanning Fluorimetry (DSF), a technique where a fluorescent marker binds to exposed hydrophobic regions of protein, was carried out for a batch of poor quality *Pa*HisG$_S$ to determine whether the inactivity was due to a stability defect. As the temperature increases during the procedure, the protein will unfold exposing hydrophobic regions. Figure 3.17 shows that compared to the original thermal denaturation of active *Pa*HisG$_S$ (T_m = 60.0 °C ± 0.1 alone, and T_m = 66.0 °C ± 0.1 with PRPP) [102], the thermal denaturation of compromised protein is considerably different with the poor batch containing a step in the thermal denaturation. This suggests that the protein may be present in two states: one folded correctly and the other not. This could account for the partial activity seen given that it is unlikely misfolded protein would be active. Determining the factor responsible for this mis-folding was not immediately forthcoming, particularly since other, completely unrelated projects were experiencing

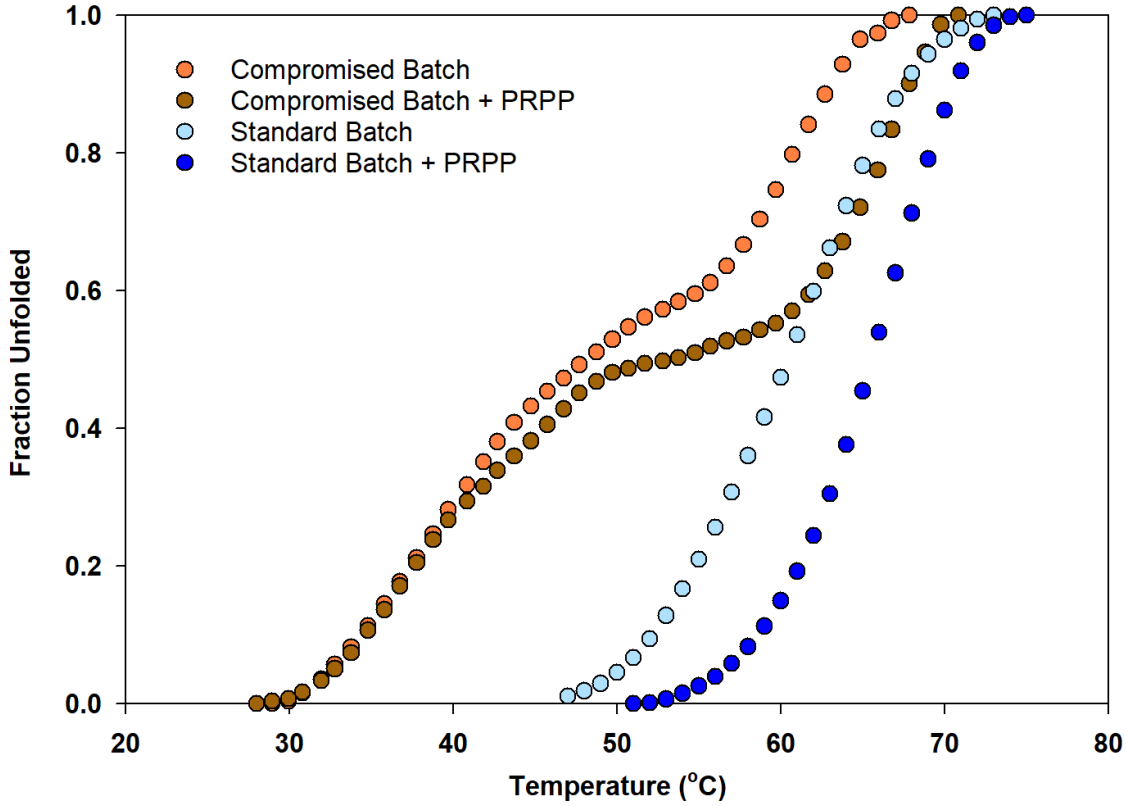


Figure 3.17: Differential scanning fluorimetry-based thermal shift assay of PRPP binding to PaHisG$_S$. Thermal denaturation of two batches of *Pa*HisG$_S$. The compromised batch is shown in orange (no ligand) and brown (PRPP ligand) and the standard batch is shown in light blue (no ligand) and brown (PRPP ligand).

similar protein purification issues.

New glycerol stocks were prepared from fresh overnight cultures but upon retesting this made no improvement to activity. All stock solutions and buffers were prepared fresh for the entirety of the procedure (including all buffers that the protein would come into contact with indirectly, for example the buffers used to strip and recharge the nickel columns) but no improvement in activity was seen. Brand new plasticware was sourced where possible or cleaned thoroughly and rinsed extensively, where not. In the event that the dialysis tubing contained detrimental contaminants, all dialysis bags were rinsed in 1 mM ethylenediaminetetraacetic acid (EDTA) before use.

Although a different HisTrap FF column was tried with no success near the beginning of troubleshooting, the alternative column had been previously used and re-charged many times in the past. It was thought that perhaps nickel was leaching out from the older column and so 1 mM EDTA was used to rinse the column and added to the purification buffers to chelate any free ions. A small improvement in *Pa*HisG$_s$activity was observed

but activity was still significant reduced. This indicated that the column may indeed be compromised. A brand new HisTrap FF column was tested and *Pa*HisG$_s$ activity was restored.

Unfortunately, despite *Pa*HisG$_s$ activity reverting back to full activity when un-complexed and un-activated, upon complexing with *Pa*HisZ, activity was lost again. Subsequently, a second factor affecting *Pa*HisZ regulation was at play and not yet accounted for. Similar factors were checked and retested as before. The main difference in purification between the two subunits is an additional volume of TEVP used, and a longer cleaving time for *Pa*HisZ than for *Pa*HisG$_S$. Many variables relating to TEVP were trialled, including purifying TEVP in fresh buffer. A batch of *Pa*HisZ without TEVP cleavage was prepared and found to have full activity. This narrowed down the potential variables causing compromised activity to a factor directly relating to TEVP.

This factor was finally determined to be the glycerol in the TEVP buffer. Correspondence with other lab groups in the building confirmed experiencing issues with glycerol, mainly over-pressuring of gel filtration columns and proteins failing to crystallise since a change of glycerol supplier from Sigma Aldrich to Scientific Laboratory Supplies. Upon purifying both TEVP and *Pa*HisZ with glycerol from Sigma Aldrich (and a glycerol free purification using phosphate buffer instead), activity was fully restored to the complex.

3.2.4 Purification of 6H*Pa*HisG$_S$ and 6H*Pa*HisZ

The 6H*Pa*HisG$_S$ (26831 Da) and 6H*Pa*HisZ (44703 Da) purified easily with no unexpected issues. Both proteins were obtained with purity >95% with yields in the region of 30 mg L^{-1} and 25 mg L^{-1} for *Pa*HisG$_S$ and *Pa*HisZ respectively. This is slightly higher than the yield of protein gained after TEVP cleavage but this is to be expected since cleavage with TEVP is not 100% efficient and some uncleaved protein is left bound to the second nickel column.

3.3 Kinetic Studies of *Pa*ATPPRT

Before inhibition kinetics of *Pa*ATPPRT were carried out, preliminary experiments regarding general kinetic characteristics were determined and confirmed to match previously published data for *Pa*ATPPRT, where applicable. Unless otherwise stated, data points are the mean of duplicates ± fitting error.

3.3.1 Determination of K_D

*Pa*HisZ (0–18 μM) was titrated into 0.67 μM *Pa*HisG$_S$ resulting in a hyperbolic relationship that was fitted to Equation 2.4 to determine the dissociation constant, K_D = 0.8 ± 0.3 µM (Figure 3.18 A). This matches the previously published K_D for *Pa*ATPPRT [20]. This K_D was used in the Morrison equation (Equation 2.4) to determine all concentrations of *Pa*ATPPRT.

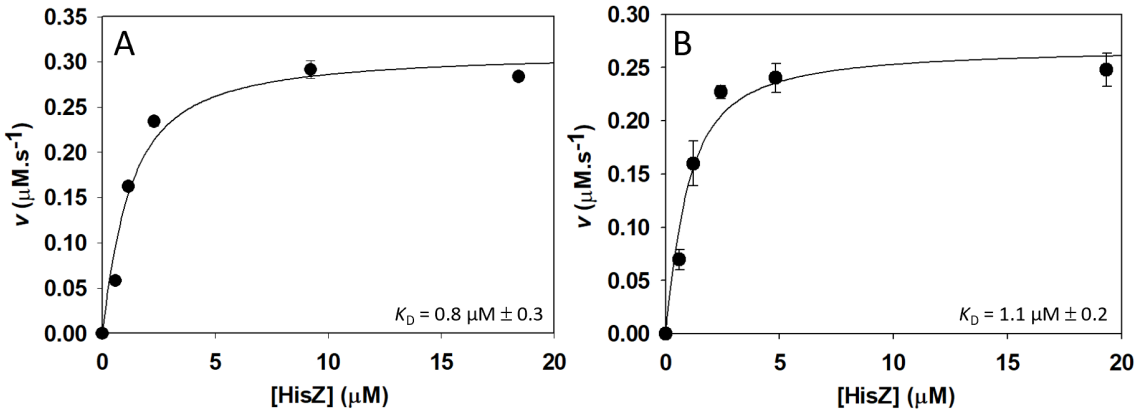


Figure 3.18: Titration of *Pa*HisZ into (A) *Pa*HisG$_S$ and (B) 6H*Pa*HisG$_S$ to determine K$_D$. The line is the best fit to Equation 2.4 and data points are the mean of duplicates ± fitting error.

3.3.2 Determination of K_M Values

To confirm the K_M values for each substrate from the current study matched those previously published [20] the rate of PRATP formation was followed while one substrate was held constant and the other substrate concentration was varied. For PRPP the ${K_M}^{PRPP}$ = 0.24 ± 0.07 mM (Figure 3.19 A). This was 2.5-fold lower than previously found for *Pa*ATPPRT (${K_M}^{PRPP}$ = 0.6 ± 0.1 mM [20]) but within error for different batches. This puts the current ${K_M}^{PRPP}$ found 5-, 16-, 21- and 13-fold higher in comparison to the ${K_M}^{PRPP}$ determined at similar temperatures for mesophilic ATPPRT's from *Mycobacterium tuberculosis*, *Campylobacter jejuni*, *Salmonella typhimurium*, and

Lactococcus lactis, respectively [58, 72, 86, 134]. Similarly for ATP, the $K_M{}^{ATP}$ = 0.68 ± 0.09 mM (Figure 3.19 B) and was 2.6-fold lower compared to that published previously for *Pa*ATPPRT ($K_M{}^{ATP}$ = 1.8 ± 0.2 mM) [20]. This places the current $K_M{}^{ATP}$ 2.6-, 7- and 6-fold higher than previously published for mesophilic ATPPRT's from *M. tuberculosis*, *C. jejuni* and *S. typhimurium*, and 4-fold lower for *L. lactis*, respectively[58, 72, 86, 134].

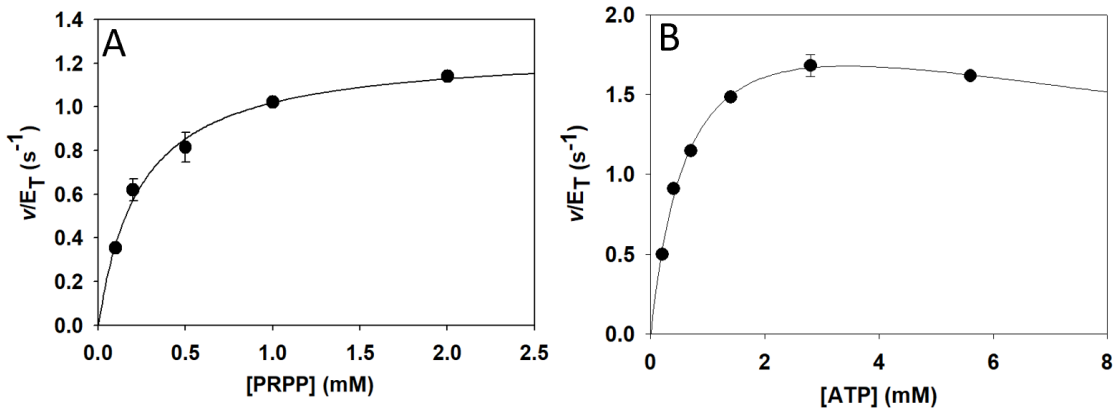


Figure 3.19: [K_M substrate saturating curves against (A) PRPP (the line is the best fit to Equation 2.5) and (B) ATP (the line is the best fit to Equation 2.6) for *Pa*ATPPRT. Data points are the mean of duplicates ± fitting error.

Surprisingly however, the current study observed substrate inhibition with a K_i = 17 ± 5 mM at ATP concentrations above 2.8 mM. Equation 2.6, which contains a term for substrate inhibition, was used to obtain the fit. This has not previously been observed in any short-form ATPPRT, including *P. arcticus*, except for ATP:Mg^{2+} complex inhibition in *S. typhimurium* (K_i = 0.8 mM) if [Mg^{2+}] was greater than 5 mM. The presence of substrate inhibition denotes that in high concentration of substrate the activity is reduced. This most-commonly observed deviation from Michaelis Menten kinetics is often an artefact of using abnormally high substrate concentrations *in vitro* [135]. Given that the K_i value is far above working concentrations, this is most likely the case here. However, substrate inhibition can, at times, prove to be of biological advantage. For example, substrate inhibition of tyrosine hydroxylase allows for a steady synthesis of the product, dopamine, regardless of fluctuations in tyrosine (substrate) concentration due to the digestion of meals [117]. Often substrate inhibition is competitive (caused by two substrate molecules binding in one binding site [136]) and this is the most likely process happening in *Pa*ATPPRT. Although the possibility of a weak allosteric regulatory substrate binding site cannot be ruled out empirically, there has been no evidence for this in previously published crystal structures with ATP bound to *Pa*ATPPRT [101].

3.3.3 Leaving the 6xHis-tag uncleaved does not alter the characteristics of *Pa*ATPPRT.

Amongst the enzymology community there is debate over whether it is essential for a histidine tag to be removed before kinetic or structural endeavours. There are many examples where leaving the 6xHis-tag on has not caused any adverse effect [59, 137, 138] but also instances where removing the 6xHis-tag has caused change [88, 139, 140]. In the present *Pa*ATPPRT study the 6xHis-tag was removed from almost all investigations so that the protein was monitored in the form closest to the one which would be present *in vivo*, while also to avoid the risk of the tag inadvertently interfering with structure or protein-protein interactions. However, a comparison of K_D, K_M and IC_{50} values were carried out between tagged and untagged protein to evaluate whether keeping the 6xHis-tags attached to the N-terminus of the protein causes any substantial differences.

Comparison of the K_D and K_M for *Pa*ATPPRT and 6H*Pa*ATPPRT

6H*Pa*HisZ (0–18 μM) was titrated into 0.17 μM 6H*Pa*HisG$_S$ which resulted in K_D = 1.1 ± 0.2 µM (Figure 3.18 B). This is within a similar region of the K_D of untagged *Pa*ATPPRT (K_D = 0.8 ± 0.3 µM) (Figure 3.18 A) and so it can be considered that the 6xHis-tag does not change the affinity of *Pa*HisG$_S$ for *Pa*HisZ.

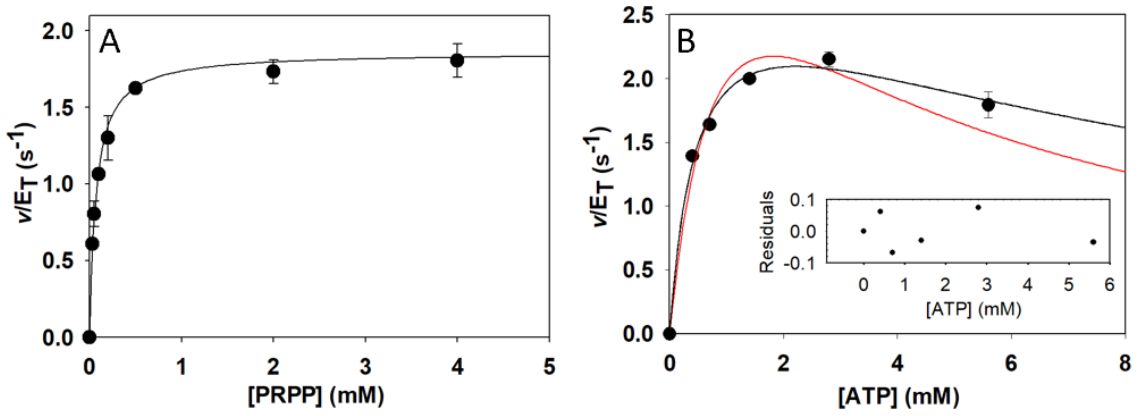


Figure 3.20: [K_M substrate saturation curves against PRPP (A) and ATP (B) for 6H*Pa*ATPPRT. In (A) the line is the best fit to Equation 2.5 and (B) to Equation 2.6 (red) and Equation 2.7 (black). Inset shows residuals of the fit to Equation 2.7. Data points are the mean of duplicates ± fitting error.

For the 6H*Pa*ATPPRT, ${K_M}^{PRPP}$ = 0.07 ± 0.008 mM (Figure 3.20 A). This was 14-fold lower than found for untagged *Pa*ATPPRT (${K_M}^{PRPP}$ = 0.24 ± 0.07 mM) in this study. This is larger than would be expected from within error only, between different batches. On

the other hand, ${K_M}^{ATP}$ = 0.5 ± 0.1 mM with K_i = 10 ± 3 mM was comparable (Figure 3.20 B) to untagged *Pa*ATPPRT in this study (${K_M}^{ATP}$ = 0.68 ± 0.09 mM with K_i = 17 ± 5 mM). This implies that substrate affinity for the enzyme is unchanged when the 6xHis-tag remains bound to the protein for ATP but substrate affinity for the enzyme increases for PRPP under these circumstances. However, the crystal structure of *Pa*ATPPRT:PRPP [101] shows the N-terminus 6xHis-tag is not in close proximity to the PRPP binding site. Perhaps this is an artefact of the purification process, since purification without TEVP cleavage is considerably shorter and hence the cold-adapted enzyme spends less time at 4 °C. A crystal structure of PRPP bound to 6H*Pa*ATPPRT could help elucidate if this is the case or a structural ramification does indeed occur when the 6xHis-tag remains on the protein.

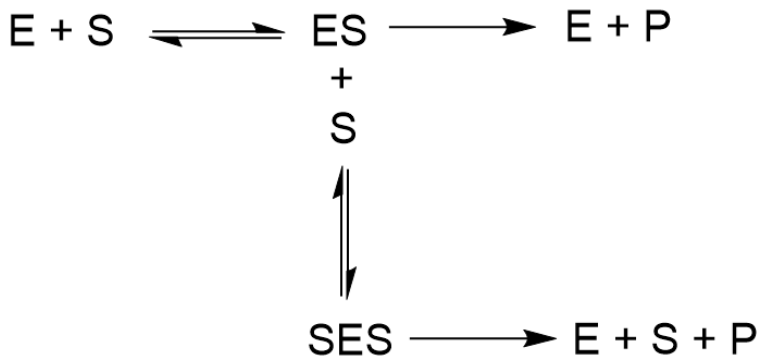


Figure 3.21: Simplified scheme of Haldane substrate inhibition for a one substrate reaction where a second substrate molecule may bind to the enzyme, forming complex SES, and resulting in inhibition.

It should also be noted that in this instance, although substrate inhibition was observed as in the cleaved protein, using the same equation to model this inhibition (Equation 2.6) did not give a convincing fit (R^2 = 0.95, Figure 3.20 red line). A much more appealing fit (R^2 = 0.99 Figure 3.20 black line) was obtained using equation 2.7 instead [117]. This equation differs from the standard substrate inhibition Equation (2.6) by including [substrate] to the power of 2. This means that as the rate tends to zero (complete inhibition), [S] becomes larger since increasingly more enzyme is unavailable due to formation of a dead-end ternary complex with two bound molecules of substrate [117]. A simplified scheme of such a reaction is shown in Figure 3.21. It is worth noting that using this equation (proposed by Haldane [141]) to describe the cleaved protein gave the same K_M and K_i as the standard substrate inhibition equation.

3.3.4 AMP is an inhibitor of *Pa*ATPPRT and *Pa*HisG$_S$

AMP is an inhibitor of *Pa*ATPPRT with IC_{50} = 333 ± 29 μM (Figure 3.22 B) and h = 0.88 ± 0.08. IC_{50} is the concentration of inhibitor (in this case AMP) required to give half the maximal activity of *Pa*ATPPRT and h is the Hill number which incorporates a term for cooperativity into Equation 2.8. If the Hill number is above 2 for complexes with four binding sites (*Pa*ATPPRT), then there is significant positive cooperativity at play. Below this value, the Hill number is of less importance, but incorporation of the term often reduces the R^2 of the fit and gives a better model [142].

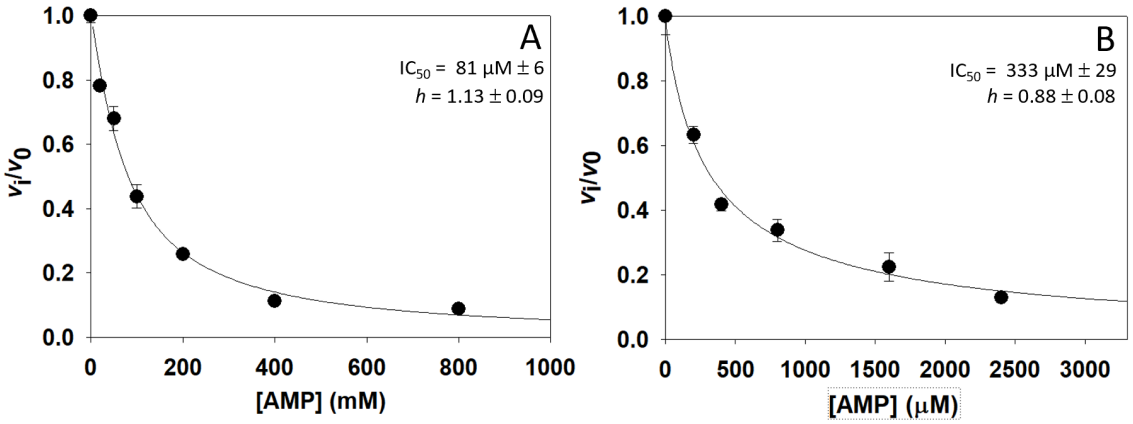


Figure 3.22: Dose-response curve for AMP with (A) *Pa*HisG$_S$ and (B) *Pa*ATPPRT. Lines are the best t to Equation 2.8. Data points are the mean of duplicates ± fitting error.

The relatively high IC_{50} value in the mid-micromolar range makes ATP a much weaker inhibitor of *Pa*ATPPRT than histidine (Section 3.4.2).*Pa*HisG$_S$ is 4x more sensitive to AMP than *Pa*ATPPRT with IC_{50} = 81 ± 6 μM and a h = 1.13 ± 0.09 (Figure 3.22 A). This suggests that the hetero-octamer desensitises the complex to inhibition by AMP. Since the AMP:ATP ratio depends on the energy status of the cell, this inhibition links the histidine pathway with the cell's metabolic status [143]. Given that the AMP inhibition is weaker for *Pa*ATPPRT than *Pa*HisG$_S$, it follows that the cell relinquishes more control to feedback inhibition by histidine for the octameric complex but is more reliant on AMP inhibition to control *Pa*HisG$_S$ activity alone. This is logical, given *Pa*HisG$_S$'s insensitivity to feedback inhibition and the requirement to conserve resources in times of low energy availability in the cell. However, there is currently no data to determine how much free *Pa*HisG$_S$ is available in the cell and therefore how much importance this inhibition has *in vivo*. AMP inhibition of ATPPRTs has been reported in HisG$_L$s (*M. tuberculosis*, *E. coli* and *C. jejuni* [86, 104, 144]) and in HisG$_S$ (*L. lactis* [91]). The AMP inhibition of *Pa*HisG$_S$ in this study was published as part of [102], a R.G da Silva group paper with kinetic data on *Pa*HisG$_S$ activation and engineering by Ms Gemma Fisher.

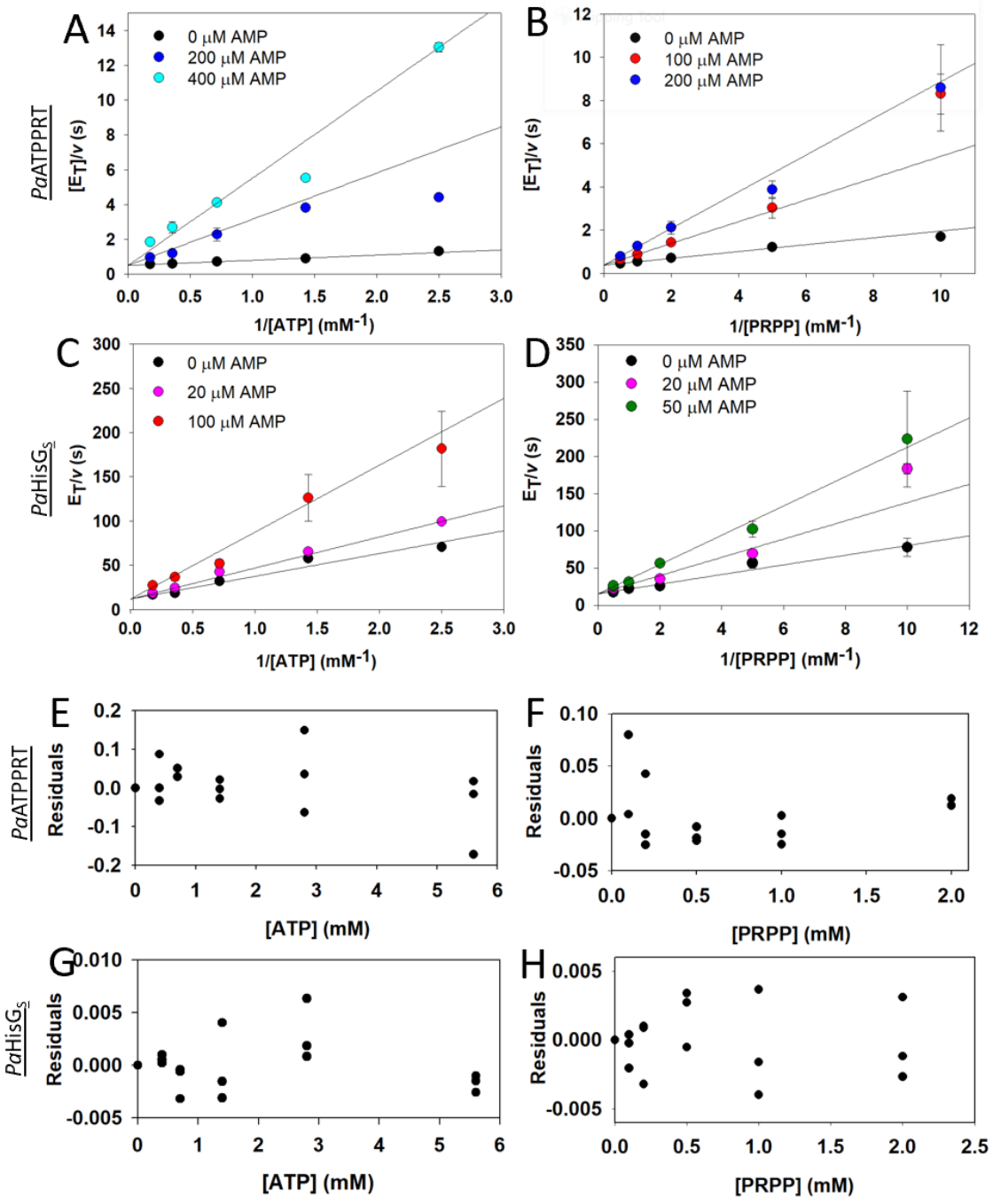


Figure 3.23: Double-reciprocals of substrate saturation curves against ATP (A and C) and PRPP (B and D) in the presence of AMP for *Pa*ATPPRT (A and B) and *Pa*HisG$_S$ (C and D). Data points are the mean of duplicates ± fitting error from individually fitted double-reciprocal plots, and lines are linear regressions of the data. Residuals from the fit of A-D to Equation 2.9. AMP inhibition against ATP (E and G) and PRPP (F and H) for *Pa*ATPPRT (E and F) and *Pa*HisG$_S$ (G and H).

AMP is a competitive inhibitor against both substrates for *Pa*ATPPRT and *Pa*HisG$_S$ (Figure 3.23), as visualised by the characteristic convergence of lines (corresponding to varying concentrations of inhibitor) on the y-axis of a double reciprocal plot. If the inhibition type had been uncompetitive, the lines on a double reciprocal plot would be parallel and for non-competitive inhibition the lines would converge to the left of the y-axis. Determined by global fitting to Equation 2.9, ATP inhibition the $K_i = 25 \pm 3$ μM against ATP and $K_i = 46 \pm 2$ μM against PRPP for *Pa*ATPPRT and $K_i = 52 \pm 8$ μM against ATP and $K_i = 25 \pm 5$ μM against PRPP for *Pa*HisG$_S$. The values for *Pa*ATPPRT are around 7-fold and 10-fold lower than those of previous characterised HisG$_L$ ATPPRTs [86, 109] and over 27-fold lower than that published for short-from ATPPRT, *L. lactis* [72]. This implies that *Pa*HisG$_S$ activity is more rigorously regulated by AMP than mesophillic ATPPRTs.

Initially it may be surprising that AMP is competitive versus both substrates, which by definition implies that AMP competes for the binding site of both PRPP and ATP. However, this study confirms the findings of the *Pa*HisG$_S$-AMP crystal structure which shows this is exactly the case (Figure 3.24). AMP sits across both the ATP and PRPP binding regions in the active site of *Pa*HisG$_S$. The 5'-PO_4^{2-} group of AMP makes similar interactions with the protein as the corresponding group in PRPP, and the adenine of AMP sits in the same site as the adenine of ATP. This is the same binding mode as found in other ATPPRTs [86, 104, 144].

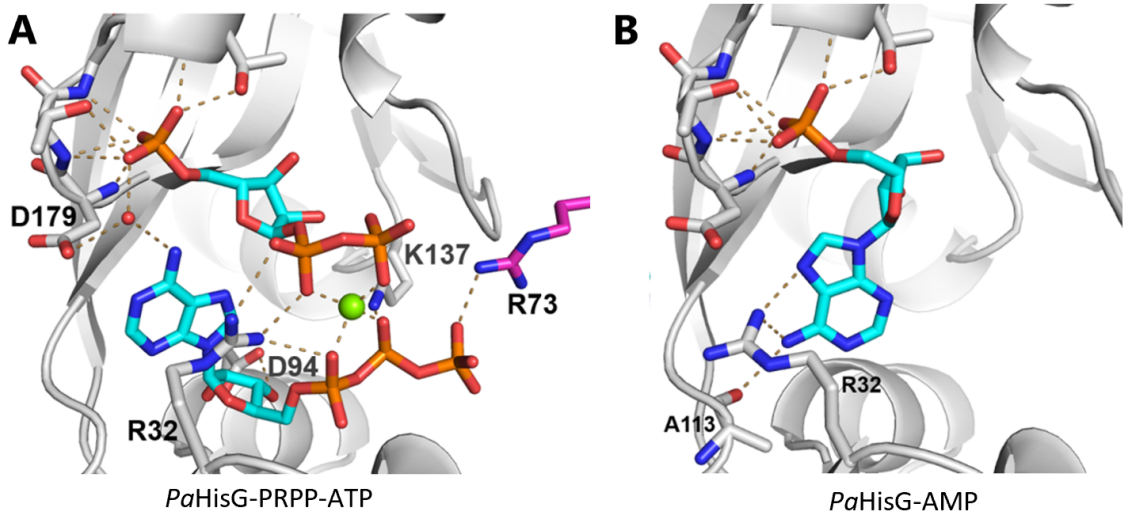


Figure 3.24: Crystal structure of (A) *Pa*HisG$_S$-PRPP-ATP and (B) *Pa*HisG$_S$-AMP binding site. The figure presents magnesium atoms (green spheres) nitrogen atoms (blue), oxygen atoms (red), phosphorus atoms (orange), ligands (stick models, carbon in cyan), protein backbone (ribbon diagram with selected main and side chain residues as stick models, carbon in grey) [101].

3.3.5 TIH is a Histidine Analogue that does not Inhibit or Activate *Pa*ATPPRT

TIH (3-(*2-thienyl)*-L-alanine) is an analogue of histidine which has a thiophene ring in place of histidine's imidazole ring and is a known allosteric activator of *M. tuberculosis* ATPPRT [92]. TIH had no effect on catalysis with *Pa*ATPPRT: no change in rate occurred when TIH was present in the assay (Figure 3.25). It is not surprising that $HisG_L$ and $HisG_S$ are activated/inhibited by different molecules since the regulatory domain of long form ATPPRTs descended from very different ancestor proteins than the regulatory subunit found in short-form ATPPRTs. However, the fact that histidine (which contains an imidazole ring that is capable of making polar contacts) is an allosteric inhibitor and TIH (with a thiophene ring which cannot make polar contacts) is not, suggests that interaction with the imidazole ring are essential for allosteric inhibition of *Pa*ATPPRT.

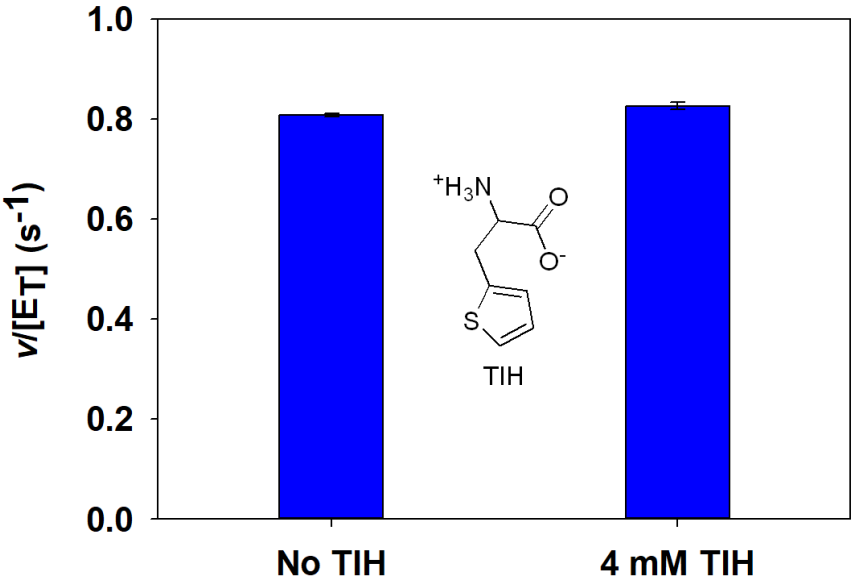


Figure 3.25: Effect of TIH on *Pa*ATPPRT rate of PRATP formation.

3.3.6 Other Effectors of *Pa*$HisG_S$

Substrate-like molecules (Figure 3.26) were tested for changes to the rate of product formation by *Pa*$HisG_S$. Ribose-5-phosphate (R5P) is a PRPP analogue lacking the phosphate but it had no effect on *Pa*$HisG_S$ activity (Figure 3.27 A). This is surprising due to R5P's similarity to the substrate. Given that the phosphate interactions with magnesium are no longer accessible, perhaps there is not sufficient structural contacts with the protein to hold R5P in the binding site long enough for sampling of conformations leading to dead-end complexes or structural changes to be observed.

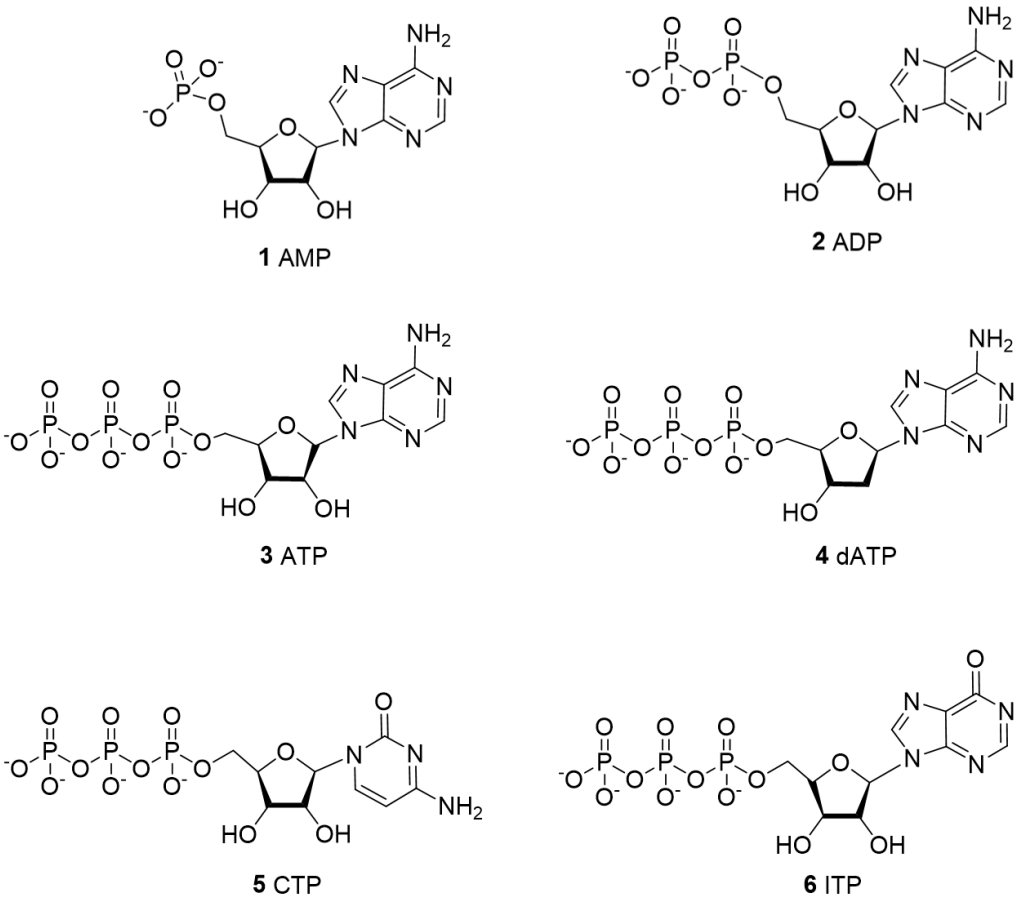


Figure 3.26: Structures of (1) AMP - adenosine monophosphate, (2) ADP -adenosine diphosphate, (3) ATP - adenosine triphosphate, (4) dATP - 2'-deoxyadenosine triphosphate, (5) CTP – cytosine triphosphate and (6) ITP – inosine triphosphate.

Given their similarity to ATP, other molecules dATP, CTP and ITP (Figure 3.26) were tested for inhibitory effects on *Pa*HisG$_S$ (Figure 3.27B). In addition, dATP was assessed as an alternative substrate for *Pa*HisG$_S$.

CTP and dATP are not inhibitors of *Pa*HisG$_S$ and dATP is not an alternative substrate as the rate was below the limit of the detector (absorbance = 0.009 and 0.0015 for both repeats). In the case of CTP it appears that the 5 membered ring-triphosphate moiety alone is not capable of stabilising the molecules in the ATP binding site to form a dead-end complex or reaction. This suggest that the contacts between the adenine ring and the enzyme are important for substrate stabilisation within the active site and triggering of the ATPPRT reaction. On the other hand, the only difference between ATP and dATP is two hydroxyl groups on the adenosine in ATP compared to only one in dATP. This implies that the hydroxyl at the 2' position on the adenosine of ATP is

that which is essential for stabilisation of the substrate within the active site (since when it is missing in dATP no reaction or inhibition occurs). This is an interesting find which is substantiated by the *Pa*HisG$_S$:PRPP:ATP structure [101] which shows Asp94 hydrogen bonds to the 2'OH (and 3'OH) of ATP. This is also true of *C. jejuni* [86] and *M. tuberculosis* [104] and so it would be interesting to determine whether mutating this residue to an amino acid incapable of interacting with the 2'OH hydroxyl would alter activity in *Pa*HisG and *Pa*ATPPRT. Since the Asp94 is a conserved residue across all 17 short-form ATPPRTs aligned in Alphey et. al [101], it is highly likely that removing this interaction would be significantly detrimental to activity.

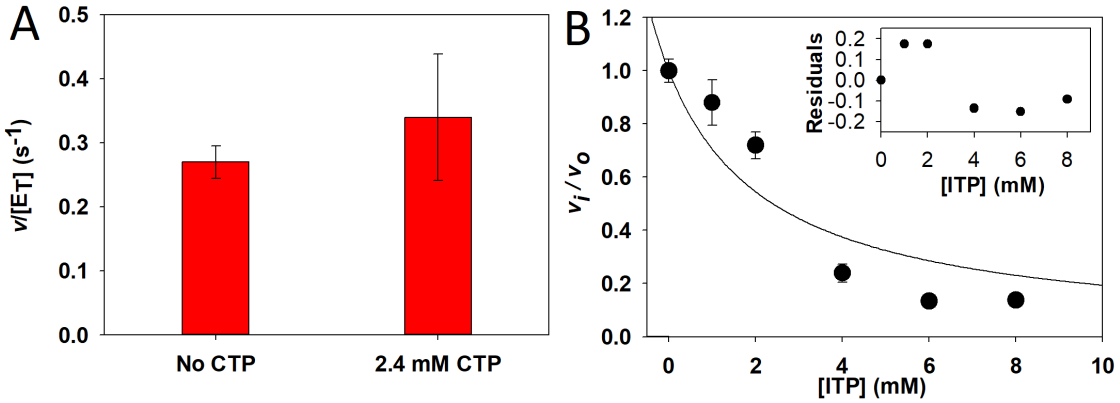


Figure 3.27: Effect of CTP (A) and ITP (B) on activity of *Pa*HisG$_S$. (B) shows a dose-response curve for ITP concentration fitted to Equation 2.8 and inset shows the residuals of the fit. Data points are the mean of duplicates ± fitting error and the inset is the residual error from the fit.

It appears that ITP is a weak inhibitor of *Pa*HisG$_S$ with IC$_{50}$ = 2400 +/- 700 μM (Figure 3.27) although the fit is not ideal (R^2 = 0.86, Figure 3.27 (B) inset) which may be related to the quality of the protein that was used for this study (see Section 3.2.3) or ITP complexing with Mg^{2+}. It could be likely that the inosine ring is able to interact with the protein enough to stabilise ITP in the ATP binding site more so than CTP. Given the minimum inhibitory effect of ITP, obtaining a crystal structure of ITP bound to *Pa*HisG$_S$ would not be simple and unlikely to be of high value.

3.3.7 ADP is an Alternative Substrate for *Pa*ATPPRT

ADP was reported as an alternate substrate for *Pa*HisG$_S$ [102] but interestingly, ADP has been reported as an inhibitor of HisG$_L$s ATPPRTs [72]. The crystal structure of *Pa*HisG$_S$-PRPP-ADP was similar to the *Pa*ATPPRT:PRPP:ADP structure [101] implying that ADP could also replace ATP as a substrate for *Pa*ATPPRT.

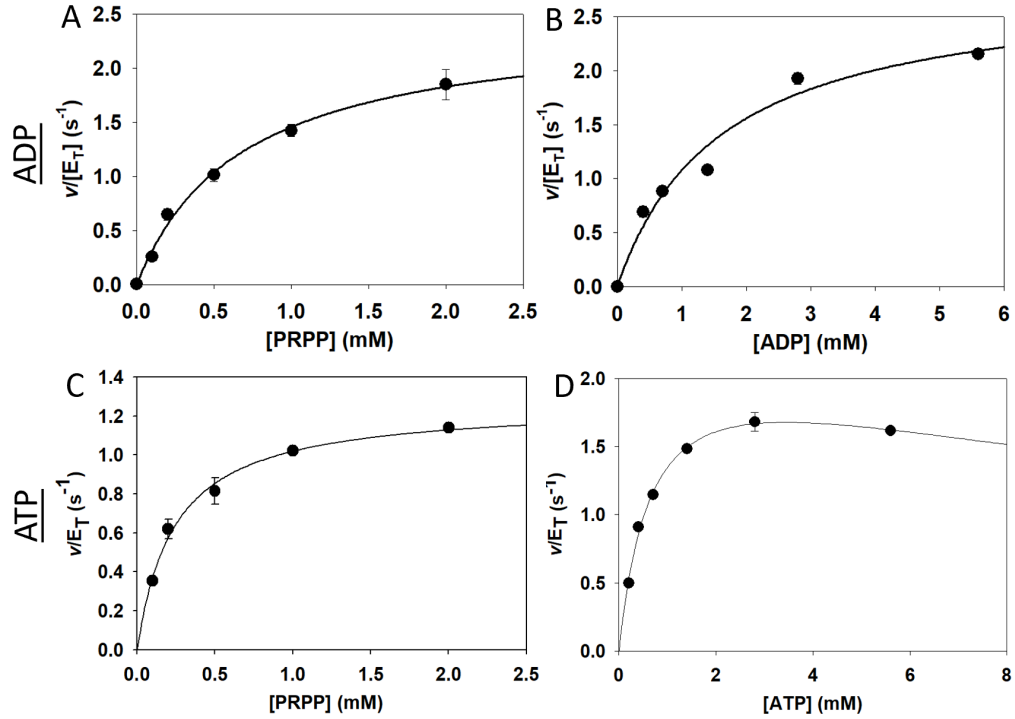


Figure 3.28: $K_m$substrate saturation curves against PRPP (A) and ADP (B) for *Pa*ATPPRT with ADP as substrate. (C) and (D) are comparison data from Figure 3.19 with ATP as substrate. The line is best fit to Equation 2.5. Data points are the mean of duplicates ± fitting error.

ADP as an alternate substrate to ATP was indeed confirmed via steady-state kinetics (Figure 3.28). The ${K_M}^{PRPP}$ = 0.7 ± 0.1 mM and ${K_M}^{ADP}$ = 0.68 ± 0.09 mM with k_{cat} = 2.6 ± 0.4 s^{-1}. These values are comparable to that found for ADP as a substrate of *Pa*HisG$_S$ (${K_M}^{PRPP}$ = 0.91 mM ± 0.08 and ${K_M}^{ADP}$ = 3.2 mM ± 0.3) [102]. Notably, no substrate inhibition was observed for ADP as the substrate.

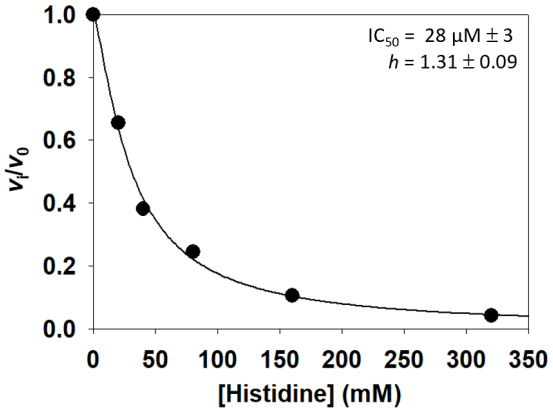


Figure 3.29: Dose-response curve for histidine with ADP as the substrate for *Pa*ATPPRT. The line is best fit to Equation 2.8. Data points are the mean of duplicates ± fitting error.

Additionally, histidine is still an inhibitor when ADP is used as the substrate. The IC_{50} = 28 ± 3 μM and h = 1.31 ± 0.09 (Figure 3.29), are effectively the same as when ATP was used as the substrate (IC_{50} = 31 ± μM 1). It appears that replacing the substrate with ADP causes no observable consequences for the formation of PRADP over PRATP by *Pa*ATPPRT or allosteric inhibition by histidine. This implies that the interaction between Arg73 of *Pa*HisG$_S$ and the γ-PO-4 group of ATP is not essential for catalysis or allosteric inhibition, since this is absent when ADP replaces ATP [102].

3.4 Kinetic Studies of *Pa*ATPPRT Inhibition by Histidine

3.4.1 Differential scanning fluorimetry thermal shift assays of *Pa*HisZ

Differential scanning fluorimetry (DSF) thermal shift assays of *Pa*HisZ in the presence of histidine were carried out to determine whether a mathematical term corresponding to histidine binding to free *Pa*HisZ would need to be included in kinetic calculations. This was determined not to be necessary, although histidine does bind to free *Pa*HisZ (Figure 3.30(A)).

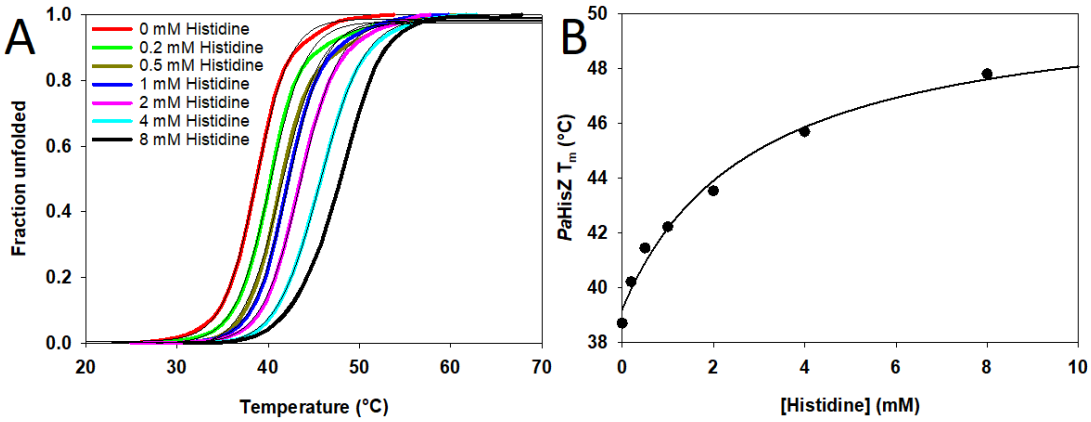


Figure 3.30: Differential scanning fluorimetry-based thermal shift assay of *Pa*HisZ in the presence of histidine. Thermal denaturation curves of *Pa*HisZ in the presence of 0 – 8 mM histidine (A). Histidine dependence on *Pa*HisZ T_m (Equation 2.1) is shown in (B). Data are the mean of quadruplicates.

The T_m of *Pa*HisZ denaturation increased upon increasing concentration of histidine. This corresponds to stabilisation of *Pa*HisZ from binding of histidine. However, the K_D of histidine binding to free *Pa*HisZ is 2.8 ± 0.8 mM (Figure 3.30 (B)). This means that if there were 2.8 mM of histidine in the assay, half of free *Pa*HisZ sites would be bound. This value is 9x higher than the maximum concentrations used for IC_{50} determination (640 μM) and 35x higher than the maximum concentration used in steady-state kinetics studies (80 μM). Therefore, any binding of histidine to free *Pa*HisZ would be negligible. Additionally, this suggests that histidine is capable of binding to *Pa*HisZ in the absence of *Pa*HisG$_S$ but at much lower affinity than when *Pa*ATPPRT is in the hetero-octameric arrangement. The very high K_D may also suggest that binding may be unspecific, and not reflect the binding mode in the hetero-octamer. It is possible that the structure of the binding site in *Pa*HisZ alone differs to that in the complex, which could result in a different K_D. Unfortunately, it is not practical to carry out a similar study on the holoenzyme as denaturation of the complex obscures the findings. Comparison of the

holoenzyme crystal structure to that of *Pa*HisZ crystallised on its own could highlight the degree of similarity of the binding sites. However, given how much higher the K_D is from working concentrations, differences in the binding sites are unlikely to be so vast as to affect calculations going forward.

3.4.2 Determination of IC_{50} for Histidine Inhibition

Histidine was confirmed as an inhibitor of *Pa*ATPPRT with an IC_{50} value of 35.5 ± 0.8 μM and h = 1.30 ± 0.04 (Figure 3.31 A). The hill number incorporates a term for cooperativity into Equation 2.8. If the Hill number is above 2 for complexes with four binding sites (*Pa*ATPPRT), then there is significant positive cooperativity at play. Below this value, the Hill number is of less importance, but incorporation of the term often reduces the R^2 of the fit and gives a better model [142].
This IC_{50} value is in agreement to both values reported for the inhibition of *M. tuberculosis* $HisG_L$ ATPPRT [58].

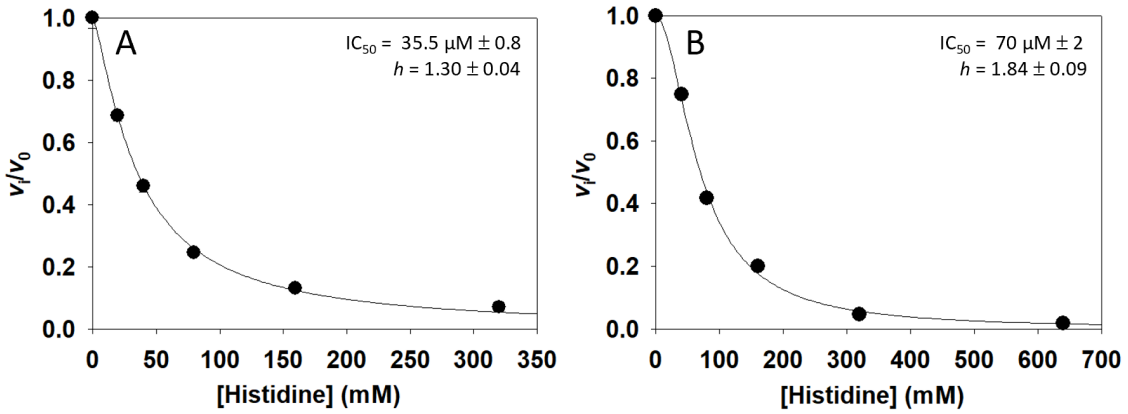


Figure 3.31: Dose-response curves for histidine concentration (A) for *Pa*ATPPRT and (B) 6H*Pa*ATPPRT. Lines are the best fit to Equation 2.8 and data points are the mean of duplicates ± fitting error.

Histidine inhibits 6H*Pa*ATPPRT with IC_{50} = 70 ± 2 μM and h = 1.84 ± 0.09, (Figure 3.31 B) comparable to untagged *Pa*ATPPRT (35.5 ± 0.8 μl). Therefore, it does not appear that removing the histidine tag has any considerable effect on histidine inhibition.

3.4.3 Determination of $K_i$values for Histidine Inhibition

For *Pa*ATPPRT, histidine is a non-competitive inhibitor versus both substrates, ATP and PRPP (Figure 3.32 and Figure 3.33), with the characteristic pattern of lines converging to the left of the y-axis on a double reciprocal plot. The K_{is}= 22 ± 6 μM and K_{ii} = 38 ± 7 μM against PRPP and K_{is} = 56 ± 25 μM and K_{ii} = 36 ± 5 μM against ATP, by

global fitting to Equation 2.10. K_{is} reports on inhibitor binding to the free enzyme, while K_{ii}, to the enzyme-substrate complex. Given that the K_{ii} and K_{is} values are similar, histidine can be considered to have almost equal affinity for the relevant enzyme forms, free *Pa*ATPPRT, *Pa*ATPPRT:PRPP, and *Pa*ATPPRT:PRPP:ATP. Histidine is also a non-competitive inhibitor of both $HisG_L$ and $HisG_S$ ATPPRTs [68, 72, 84, 86]. Of the $HisG_L$ ATPPRTs characterised, only the *M. tuberculosis* enzyme differs with inhibition uncompetitive against ATP, although the K_i's were still comparable [58]. K_i's were also comparable for ATPPRTs from *C. jejuni* [86], and *L. lactis* [72] but almost 10x lower than the K_i's stated for *S. typhimurium* [134] and *Thermotoga maritima* [84] ATPPRTs.

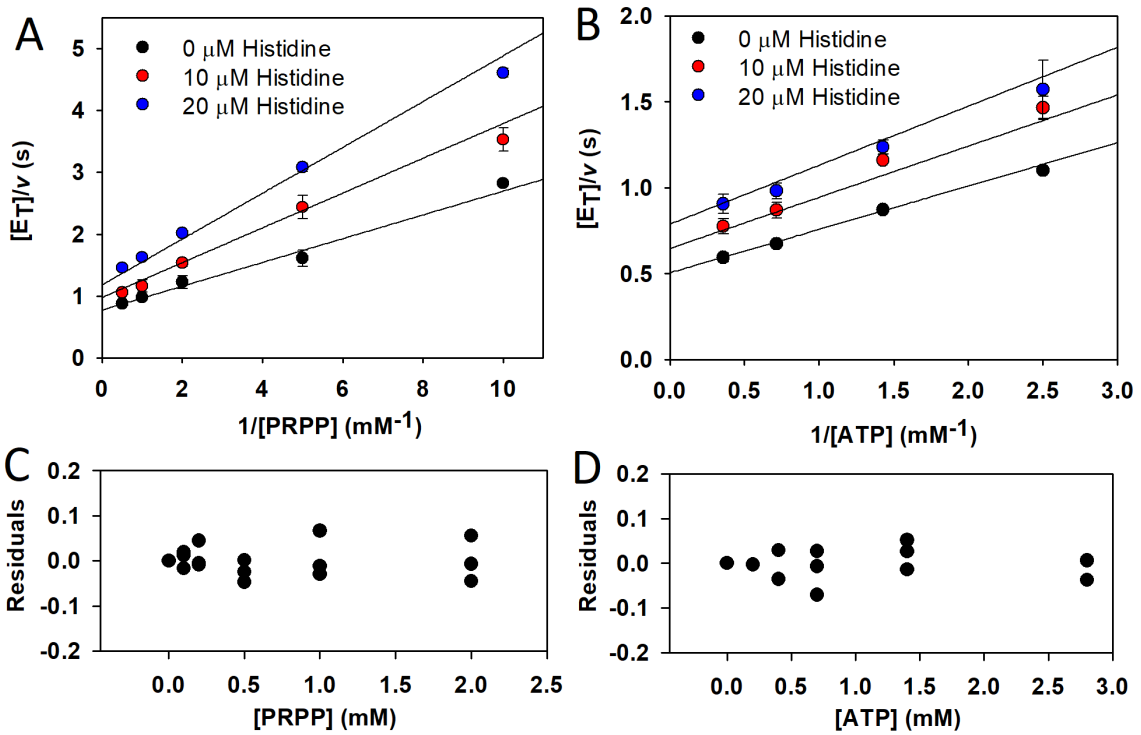


Figure 3.32: *Pa*ATPPRT inhibition by histidine. Double-reciprocals of substrate saturation curves in the presence of histidine against PRPP (A) and ATP (B) by global fitting to Equation 2.10. Residuals of the model fitting are shown for PRPP (C) and ATP (D). Data points are the mean of duplicates ± fitting error from individually fitted double-reciprocal plots, and lines are linear regressions of the data.

Replots of histidine inhibition data. Slope and intercept replots of the individually fitted double-reciprocal data for PaATPPRT inhibition by histidine against PRPP and ATP. Data points are value ± fitting error from individually fitted double-reciprocal plots, and lines are linear regressions of the data.

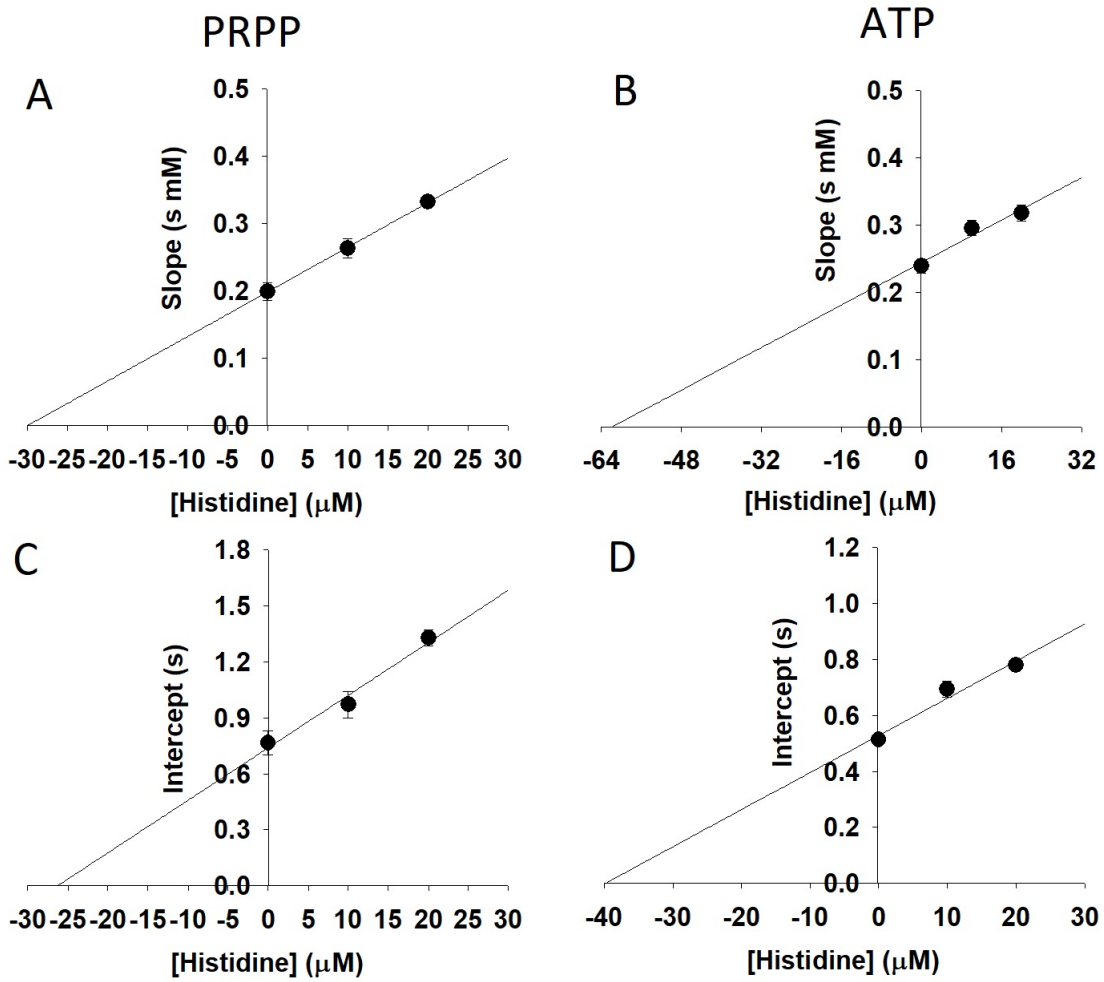


Figure 3.33: Replots of histidine inhibition data. Slope and intercept replots of the individually fitted double-reciprocal data for *Pa*ATPPRT inhibition by histidine against PRPP and ATP. Data points are value ± fitting error from individually fitted double-reciprocal plots and lines are linear regressions of the data.

3.4.4 Effect of Histidine on the Approach to Steady State.

Approach to steady state, or 'pre-steady state' kinetics refers to the period before the steady state is reached. Studying the formation and consumption of products and reactants during this period can inform on binding order and the rate limiting step of the reaction. *Pa*ATPPRT follows an ordered mechanism with PRPP binding to the enzyme first [102]. This is the opposite order to other ATPPRTs characterised [69, 75] but product release follows the same pattern as both long and short-form ATPPRTs previously published with PP_i leaving first, followed by PRATP (Figure 3.34).

The rate limiting step for *Pa*ATPPRT catalysis is product release (as determined by the presence of an initially large rate or 'burst' in previous pre-steady kinetic experiments)

[102]. This contrasts with *Pa*HisG$_S$ alone, where the rate-limiting step is likely chemistry. It can therefore be said that the rate limiting step of *Pa*ATPPRT is determined by the activation, or not, of *Pa*HisG$_S$. The effect of histidine on the approach to steady state is hence of interest.

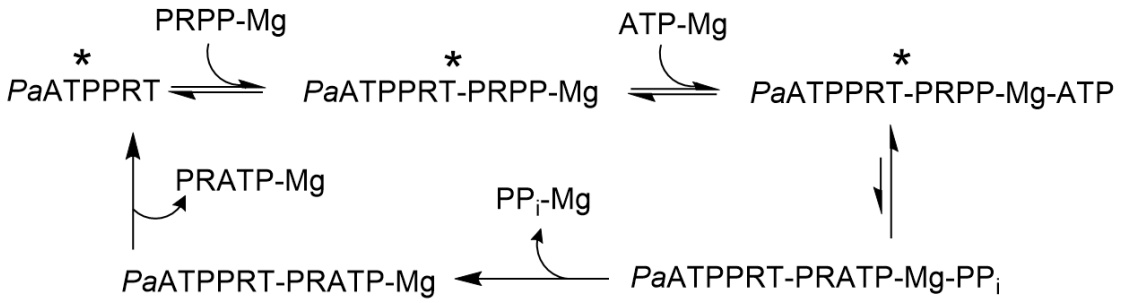


Figure 3.34: Kinetic mechanism of *Pa*ATPPRT is sequential and ordered with ATP binding after PRPP and PRATP leaving after PPi. Histidine can bind to all three forms of the enzymes depicted by an asterisk.

When histidine in not present in the assay and ATP is rapidly mixed with *Pa*ATPPRT pre-incubated with PRPP, an initially large rate or 'burst' is observed before the steady-state (Figure 3.35, red line). The amplitude of this burst phase is 3.11 ± 0.01 μM with an uninhibited steady state rate of 16.36 ± 0.01 μM s^{-1}. This burst amplitude remains the same, within experimental error, at 3.49 ± 0.01 μM when histidine is present at 640 μM and pre-incubated with ATP upon mixing with *Pa*ATPPRT and PRPP (Figure 3.35, green line). The steady state rate decreases, as expected by inhibition, to 7.56 ± 0.01 μM s^{-1} for this reaction. However, when histidine is preincubated with *Pa*ATPPRT and PRPP before rapidly mixing with ATP (Figure 3.35, blue line), the amplitude of the burst phase is reduced to 0.53 ± 0.01 μM despite the steady state rate of the inhibited reaction being the same as before, at 7.83 ± 0.01 μM s^{-1}. Since the histidine is all in one syringe, a small lag phase due to mixing of histidine would be expected and is observed in the latter condition. An additional control with the histidine distributed equally between both syringes could be beneficial to confirm this effect. Taken together, these data implies that the rate of ATP binding to the *Pa*ATPPRT:PRPP complex and product formation is faster than histidine binding to the same complex and inhibition occurring. Since the burst is significantly reduced when histidine is pre-incubated with the *Pa*ATPPRT:PRPP complex, histidine must be able to bind to the binary complex. Furthermore, it must also be able to form an inhibited complex in the absence of ATP since the burst amplitude is considerably smaller and product release is therefore no longer rate limiting. In the only other approach to steady state study carried out for an ATPPRT, *M. tuberculosis* ATPPRT (which follows an ordered reaction with ATP binding first) the burst amplitude decreased when ATP and histidine were rapidly mixed [58]. This contrasts with the current findings for *Pa*ATPPRT. The error presented for experiments using the Stopped

Flow spectrophotometer are from technical replicates only. Due to the large amount of protein required, a biological replicate was not carried out specifically for each experiment, although many batches of protein were used across the body of the work. Each new batch of protein used was confirmed to have similar k_{cat} and K_m.

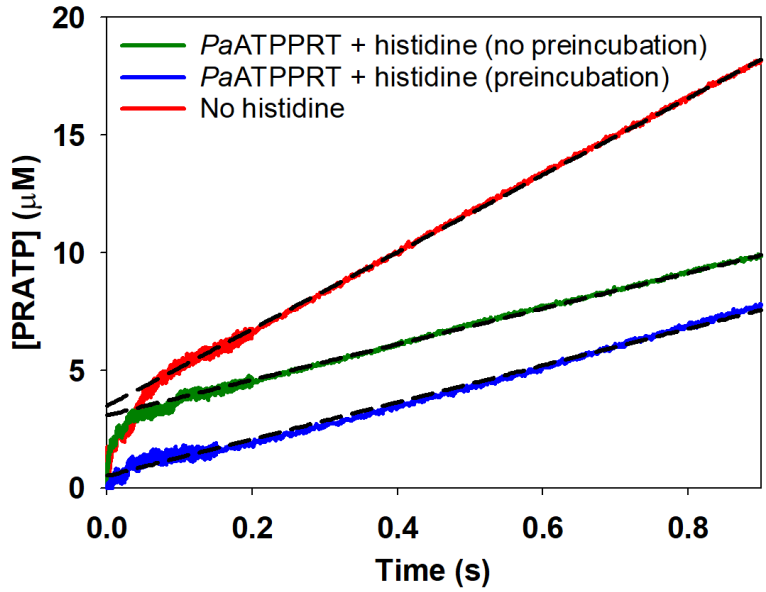


Figure 3.35: Rapid kinetics of histidine inhibition. Rapid mixing of ATP with *Pa*ATPPRT and PRPP (red line), ATP and histidine with *Pa*ATPPRT and PRPP (green line), and ATP with PaATPPRT, PRPP and histidine (blue line) Dashed lines are linear regression. Final concentrations in the mixing chamber were 19 μM *Pa*ATPPRT, 0 or 30 μM histidine, 2 mM PRPP, 3.5 mM ATP.

3.4.5 Inhibition by Histidine does not Involve Release of Non-Activated *Pa*HisG$_S$

Since *Pa*HisG$_S$ has considerably reduced activity without *Pa*HisZ, histidine's mode of inhibition on *Pa*ATPPRT could simply be to break apart the complex. This would release free *Pa*HisG$_S$, which being insensitive to histidine, would be capable of utilising substrate present in the assay and forming PRATP. If histidine were breaking up the *Pa*ATPPRT complex to give free, non-activated *Pa*HisG$_S$, then the rate of PRATP formation of inhibited *Pa*ATPPRT would be the same as the rate of *Pa*HisG$_S$ alone. This can be tested in a kinetics experiment but owing to the large discrepancy in enzyme concentration required for *Pa*HisG$_S$ activity compared to *Pa*ATPPRT, a conventional steady state assay in a UV-VIS spectrophotometer and manual mixing would deviate from linearity too soon. Therefore, a stopped flow spectrophotometer was used to catch the initial seconds of the reaction, when PRATP formation is still linear even at relatively high enzyme concentrations. When assayed, the rate of PRATP formation was lower for histidine

inhibition of 2 μM *Pa*ATPPRT (purple) than for 2 μM *Pa*HisG$_S$ (yellow) (Figure 3.36). This was confirmed to be a statistically significant result by Student's t-test with $p < 0.0031$. Therefore, histidine does not inhibit the complex merely by setting *Pa*HisG$_S$ free from the hetero-octamer. The rate of reaction upon histidine inhibition results in a decrease in rate to that of between 1 μM and 2 μM *Pa*HisG$_S$ activity alone and so histidine must deregulate activity through a different mechanism.

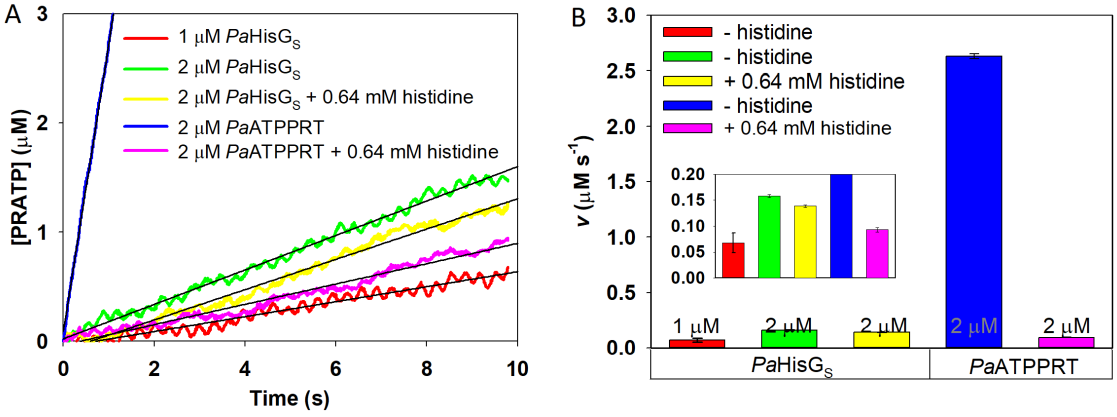


Figure 3.36: Rapid kinetics of steady state histidine inhibition. (A) Time course of *Pa*HisG$_S$ and *Pa*ATPPRT reaction with and without 0.64 mM histidine. Lines are linear regression. (B) Initial rates of PRATP formation extracted from (A). Inset is enlarged view of (B). Final concentrations in the mixing chamber were 1 or 2 μM *Pa*ATPPRT or *Pa*HisG$_S$, 0 or 640 μM histidine, 2 mM PRPP, 5.6 mM ATP.

The error presented for experiments using the Stopped Flow spectrophotometer are from technical replicates only. Due to the large amount of protein required, a biological replicate was not carried out specifically for each experiment, although many batches of protein were used across the body of the work. Each new batch of protein used was confirmed to have similar k_{cat} and K_m.

3.4.6 *Pa*HisG$_S$ Mutant Sensitivity to Histidine in *Pa*ATPPRT Complex

Previous work from our lab [102], [101] show there is very little variation in structure between *Pa*HisG$_S$ and *Pa*ATPPRT despite the presence of *Pa*HisZ leading to a large activation of *Pa*HisG$_S$. There is an observed tightening of the overall *Pa*HisG$_S$ dimer upon *Pa*HisZ binding which increases the cross-dimer contacts of some side chains, but the most striking difference between *Pa*ATPPRT and *Pa*HisG$_S$ lies in the deviations that occur between the PRPP-, PRPP-ATP- and PRART-bound structures of *Pa*ATPPRT. The position of Arg56 which is 12 Å apart from the PRPP β-PO$_4$ in the binary complex shifts considerably to form a salt bridge with the PP$_i$ moiety of PRPP after the ternary complex forms (Figure 3.37). This residue which traverses towards the active site only upon ATP binding, sits in the middle of a flexible loop and is responsible for leaving

group stabilisation at the transition state. Arg56 is highly conserved across the 18 species of ATPPRT aligned in Alphey *et. al.* with only one species, *T. maritima* containing a deletion at that site instead [101]. After the reaction has occurred, Arg56 moves again, to reside 6 Å away from its ternary complex location. This substantial movement of Arg56 is probably less favoured in *Pa*HisG$_S$, the non-activated form of the enzyme. Based on structure, there is clear support for leaving group stabilisation as a key component of *Pa*HisG$_S$ activation by *Pa*HisZ with the catalytically favoured position of Arg56 sampled more often in *Pa*ATPPRT than in *Pa*HisG$_S$.

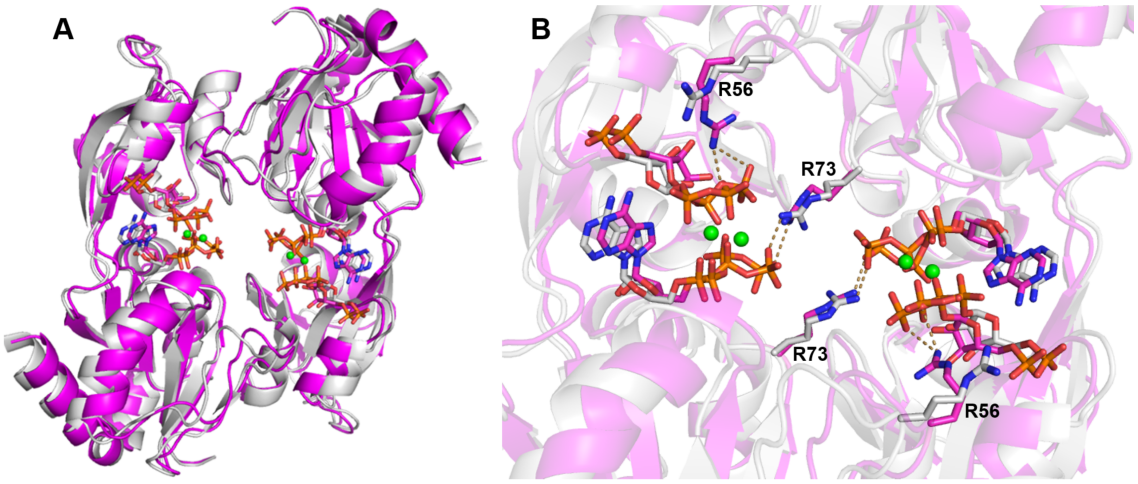


Figure 3.37: *Pa*HisZ-induced active site preorganization in *Pa*ATPPRT: (A) overlay of PRPP-ATP-bound *Pa*ATPPRT (magenta) and *Pa*HisG$_S$ (gray) dimers; (B) active site close-up of the superimposed dimers showing cross-subunit contacts. magnesium ions are shown as green spheres, while substrates and Arg56 and Arg73 side chains are represented as stick models with nitrogen in blue, oxygen in red, phosphorus in orange, and carbon in magenta (PaATPPRT) and gray (*Pa*HisG$_S$). Polar interactions are depicted by dashed lines [101]

Previously [102], the R56A-*Pa*HisG$_S$mutant was shown to have 86-fold lower activity than WT- *Pa*HisG$_S$, but only 6-fold less activity when in complex with *Pa*HisZ. This suggests that other conformations with leaving group stabilisation must still be sampled without an Arg56 salt bridge. Additionally, the R56A-*Pa*HisG$_S$ mutant complexed with *Pa*HisZ was proven, in this work, to retain histidine sensitivity (Figure 3.38). Hereby excluding the Arg56 residue from a role in allosteric inhibition in addition to allosteric activation.

As an additional area of interest in Ms Gemma Fisher's work, she mutated C115 in *Pa*HisG$_S$ based on a hypothesis that it maybe acting as a general base during the *Pa*HisG$_S$ reaction. Unpublished data by Ms Fisher showed by NMR and steady state kinetics that the C115S mutant lost all activity when *Pa*HisG$_S$ is not activated, but regains full activity when activated by *Pa*HisZ (not shown). In this current work, the C115S mutant was tested to determine if it was still sensitive to histidine. The C115S mutant did retain

sensitivity to histidine (Figure 3.38) and is a puzzling area of ongoing investigation.

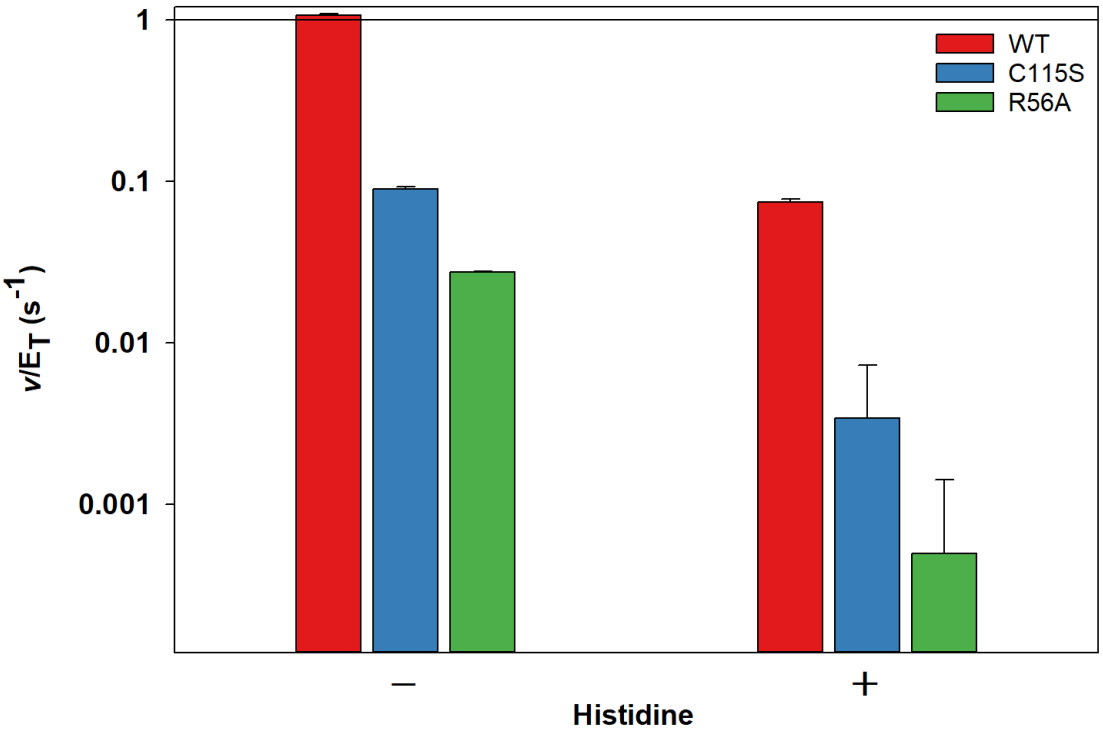


Figure 3.38: Sensitivity of WT *Pa*HisG$_S$ (red), C115S-*Pa*HisG$_S$ mutant (blue) and R56A-*Pa*HisG$_S$ (green) to 640 μM histidine. Presented on a logarithmic scale.

3.5 Crystallographic Studies of *Pa*ATPPRT Inhibition by Histidine

In order to fully understand the mode of allosteric inhibition of *Pa*ATPPRT by histidine, a plausible crystal structure is sought. Based on the kinetic data suggesting that histidine can bind to the binary and ternary complexes, both substrates were included in the crystallisation.

3.5.1 Crystallisation Trials, X-ray Data Collection and Processing of *Pa*ATPPRT:Histidine

Crystallisation conditions for *Pa*ATPPRT:ATP:PRPP:histidine were repeated as previously published [101] using 11% PEG 3350, 0.1 M bicine (pH 8.5), 0.15 M $SrCl_2$, 0.15 M KBr, and 2% 1,6-hexanediol at 4 °C. *Pa*HisG$_S$ and *Pa*HisZ were mixed in 1:1 molar ratio, and the buffer was exchanged with 0.02 M Tris (pH 7.0), 0.05 M KCl, 0.01 M $MgCl_2$, 2 mM DTT, 0.5 mM histidine, and 10 mM ATP. Repeating these conditions yielded a few, small, cube-shaped crystals no larger than 0.05 × 0.05 × 0.03 mm^3, as before, per batch. Despite histidine being present in the crystallisation buffer, no histidine bound structure has ever been found with these conditions. ATP, PRPP and histidine were soaked into the crystals.

Soaking trails of the crystal with substrate and histidine were carried out by soaking the crystal in each of the substrates- and histidine- containing buffers for 3–60 min. From the initial trial, only the fastest soaking (10 minutes Soaking Buffer 2 and 20 s Soaking Buffer 3) diffracted successfully in-house. This crystal was retained, and further data collected at the Diamond Light Source in order to improve resolution and data quality.

Initially, the successfully soaked crystal was processed as described in Section 2.7.3 at 2.63 Å and space group P 1 2_1 1. Density corresponding to PRPP and histidine were present with some partial density for adenine. However, although through repeated rounds of refinement, the R_{factor} = 0.200 and R_{free} = 0.282 were within acceptable limits, the structure was plagued with numerous poor rotamers, strained bond angles and a dubious amount of Ramachandran outliers (Figure 3.39). Additionally, there was not enough density available in many of the side chains to pinpoint precisely their position. To rectify this, side chains could be cut, but given the magnitude of the problem a vast number would need to be removed, decreasing the overall quality of the output. Therefore, the refinement of this crystal was abandoned at this point. Fortuitously, a separate crystal from the same growth condition had been processed in parallel by our

collaborator Dr Magnus Alphey. This crystal had similar setbacks when refining but to a much lesser extent and the structure was fully resolved at 2.65 Å. The structure was deposited in the Protein Data Bank (PDB), entry 6R02. Table 3.2 outlines the crystallographic parameters and statistics of the final structure. Going forward, these data were used for all further analysis.

All-Atom Contacts	Clashscore, all atoms:	10.9		97th percentile (N=226, 2.63 Å ± 0.25 Å)
	Clashscore is the number of serious steric overlaps (<0.4 Å) per 1000 atoms.			
Protein Geometry	Poor rotamers	194	9.80%	Goal: 0.3%
	Favoured rotamer	1533	77.42%	Goal: >98%
	Ramachandran outliers	42	1.81%	Goal: <0.05%
	Ramachandran favoured	2105	90.73%	Goal: >98%
	MolProbity score	2.81	63rd percentile (N=6042, 2.63 Å ± 0.25 Å)	
	Cβ deviations >0.25 Å	20	0.92%	Goal: 0
	Bad bonds	27/18564	0.15%	Goal: 0%
	Bad angles	140/25165	0.56%	Goal: 0.1%
Peptide Omegas	Cis prolines	0/84	0.00%	Expected: ≤ 1 per chain, or 5%
	Cos non-prolines	1/2251	0.04%	Goal: 0.05%
	Twisted peptides	3/2335	0.13%	Goal: 0

Figure 3.39: Molprobity output after multiple rounds of refinement for the initial crystal processed. Red indicates the required standard has not been achieved. Green indicates the value obtained conforms to the required standard. Yellow indicates borderline agreement with the required standard.[145]

Table 3.2: Crystallographic data and refinement statistics for *Pa*ATPPRT:PRPP:histidine

PDB entry	6R02
Structure	*Pa*HisG$_S$Z + histidine + PRPP
Beamline	I24
Wavelength (Å)	0.9686
Resolution (Å)	77.31 – 2.65
Space group	$P2_1$
No. molecules in a.u.	4 *Pa*HisZ + 4 *Pa*HisG$_S$
Total No. reflections	254279
No. unique reflections	75715
Multiplicity	3.4 (3.4)
Completeness (%)	99.9 (99.8)
<I/sigI>	6.5 (0.8)
R_{merge}[a]	0.072 (1.247)
CC(1/2)	0.998 (0.456)
Wilson B (Å^2)	60.6
Unit cell parameters	
a (Å)	93.3
b (Å)	147.3
c (Å)	99.3
α (°)	90.0
β (°)	103.1
γ (°)	90.0
Refinement	
No. reflections (working set)	71962
No. reflections (R_{free} test set)	3726
R_{work}/R_{free}[b/c]	0.24 / 0.27
r.m.s.d bond distances (Å)	0.004
r.m.s.d. bond angles (°)	1.286
Protein atoms	17880
Ligands (histidine+PRPP)	4 + 4
Waters	54
Average B factor (Å^2)	
Protein	85.6
Ligands (PRPP + histidine)	136.6 + 74.6
Waters	60.2
Ramachandran plot	
Favoured (%)	97.0
Outliers	0

[a] $R_{merge} = \sum|I_{(hkl)} - I| \times 100/\sum|I_{(hkl)}|$, where the average intensity I is taken over all symmetry equivalent measurements and $I_{(hkl)}$ is the measured intensity for a given observation. [b] $R_{work} = \sum_{hkl}||F_o(hkl)| - |F_c(hkl)||/\sum_{hkl}|F_o(hkl)|$. [c] $R_{free} = R_{factor}$ for a test-set of reflections (5%).

Analysis of the *Pa*ATPPRT:PRPP:histidine crystal structure.

3.5.2 Histidine Binds in the Site Conserved in HisRS

The *Pa*ATPPRT:PRPP:histidine structure crystallised in the $P2_1$ space group with the entire hetero-octameric complex within the asymmetric unit. As with the previous *Pa*ATPPRT structures [20, 101, 102], *Pa*HisZ subunits sit crossed in the middle of the octameric complex with *Pa*HisG$_S$ dimers flanking top and bottom (Figure 3.40A). Density corresponding to PRPP was found in all four *Pa*HisG$_S$ binding sites as in previous structures [101]. When PRPP is bound to *Pa*ATPPRT the substrate is secured to the active site via main and side chain interactions between the PRPP binding loop and the phosphate of PRPP. There is also a hydrogen bond (H-bond) between the substrate 2-OH group and side chain residue Asp176 that stops PRPP from blocking the ATP binding site. This is similar to the PRPP bound structure of *L. lactis* [70] and *Pa*HisG$_S$ alone, but not *C. jejuni* [86].

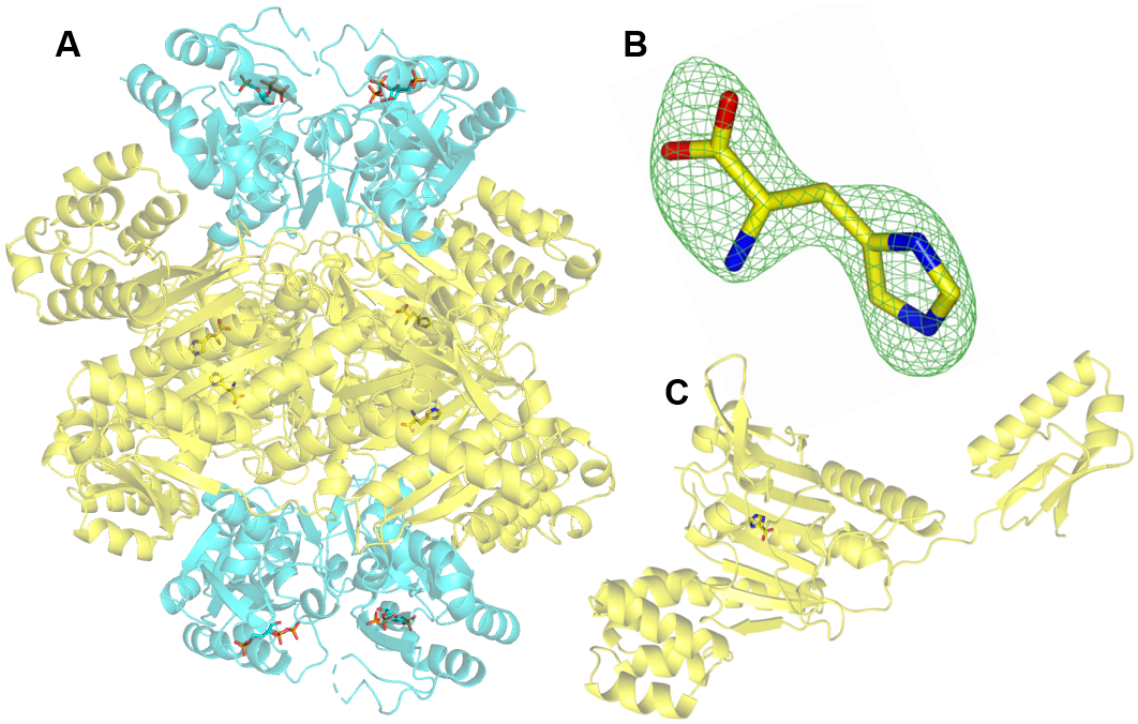


Figure 3.40: Crystal structure of *Pa*ATPPRT:PRPP:histidine. (A) Ribbon diagram of the *Pa*ATPPRT hetero-octamer found in the asymmetric unit. (B) Omit map showing electron density for histidine. (C) Ribbon diagram of the *Pa*HisZ subunit with the location of the bound histidine. *Pa*HisG$_S$ subunits are in cyan, *Pa*HisZ in yellow. Ligands are shown in stick models with carbon atoms matching the colour of the subunits to which they are bound.

Here, for the first time, clear electron density for a histidine molecule (Figure 3.40 B) has been reported in the binding site analogous to HisRS, along the central antiparallel β-sheet of each molecule of *Pa*HisZ (Figure 3.40C). No density for histidine was found at the interface of *Pa*HisG$_S$ and *Pa*HisZ subunits, contradicting the only previously published structure of a short-form ATPPRT:histidine complex (*T. maritima*) which depicted two histidine binding sites at the catalytic and regulatory sub-unit interface [84]. However, this interface resides over 18 Å away from the histidine binding site occupied in the *Pa*ATPPRT structure. Figure 3.42 highlights the discrepancy with the *T. maritima* binding sites (blue boxes) but the striking agreement with the histidine binding site of *Burkholderia thailandensis* HisRS (red box). Given that HisZ and HisRS share a common protein ancestor, this finding is reasonable and confirms previous speculation [70].

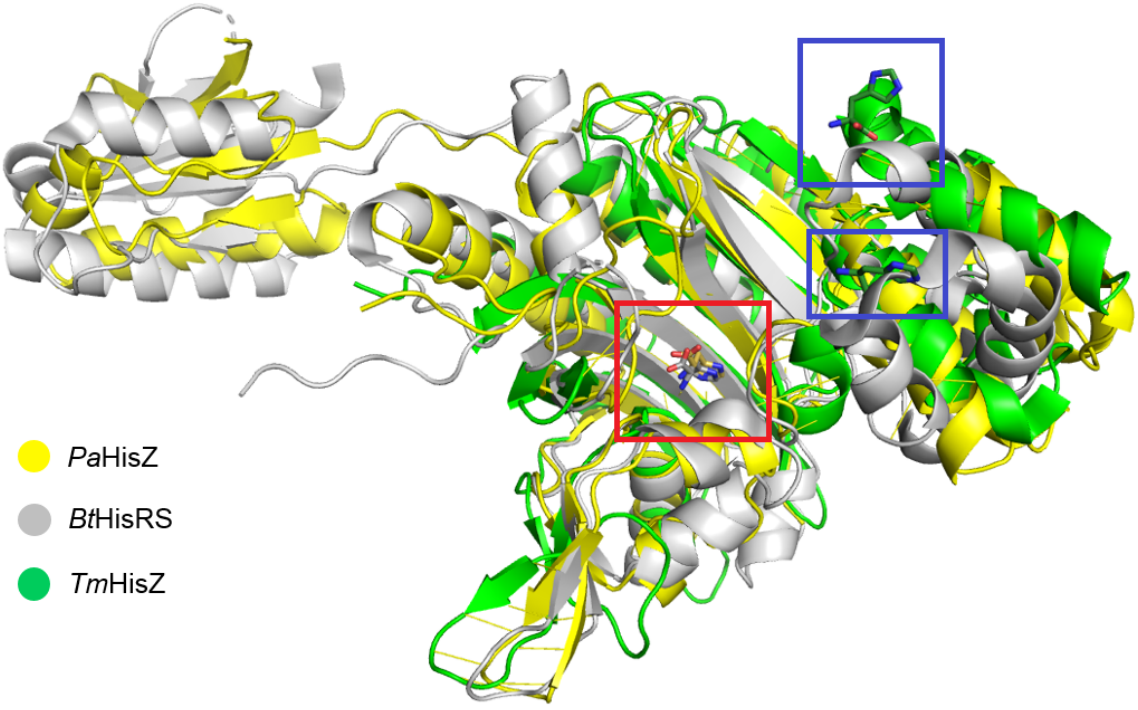


Figure 3.41: Overlay of histidine-bound *Pa*HisZ, *T. maritima* HisZ, and *B. thailandensis* HisRS structures. Histidine molecules are shown in stick models. *T. maritima* histidine binding sites are highlighted with blue boxes and the *Pa*HisZ and *B. thailandensis* HisRS histidine binding sites are shown in a red box.

3.5.3 The Histidine Binding Site

Histidine is held within the *Pa*HisZ binding site by an extensive array of polar interactions to protein side chain residues (Figure 3.42). The α-COO- group of histidine contacts the δ-NH and ω-NH2 groups of Arg284 and the γ-NH2 group of Gln118. Histidine α-NH_2 interacts via hydrogen bonds with Tyr265 4-OH group and Ser308 and Thr78 β-OH groups, and via a salt-bridge with Asp76 β-COO- group. The imidazole

ring of the inhibitor is anchored to the β-strand via an interaction between its τ-NH and Glu122 γ-COO-, and has further contact with a flexible loop that also includes Tyr265 via an interaction between its π-N and His266 τ-NH. Given that histidine is an allosteric inhibitor of *Pa*ATPPRT and TIH is not, these imidazole interactions may be of vital importance to the mode of allosteric inhibition since in their absence, no inhibition occurs.

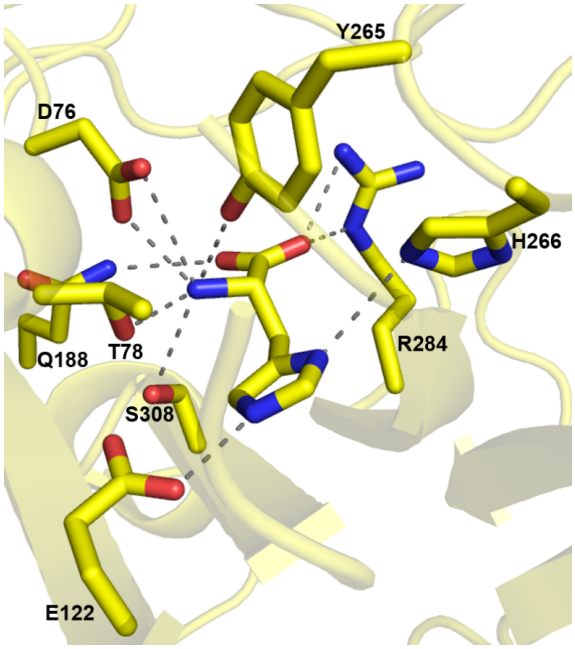


Figure 3.42: Enlarged view of the histidine binding site in *Pa*HisZ. The figure shows histidine and relevant side chains in stick models, carbon (yellow), nitrogen atoms (blue), oxygen atoms (red), protein backbone (ribbon diagram) and polar contacts depicted by dashed lines.

3.5.4 Structural Basis of *Pa*ATPPRT Inhibition and Consequences of Allosteric Regulation

The allosteric activation of *Pa*HisG$_S$ by *Pa*HisZ (where the dimer contracts, allowing the *Pa*HisG$_S$ loop movement and stabilisation of the leaving group transition state), only occurs after ATP binding [102]. Therefore, the *Pa*ATPPRT:PRPP complex is in a non-activated state. When the structure of *Pa*ATPPRT:PRPP:histidine is overlaid with this structure the Cα root-mean-square deviations (rmsd) was 0.40 rmsd (Figure 3.43A). Similarly, when the *Pa*HisZ tetramers are overlaid the rsmd is 0.38 (Figure 3.43B) and the *Pa*HisG$_S$ dimers have rmsd of 0.27 (Figure 3.43C). The overlaid structures of *Pa*ATPPRT:PRPP:ATP and *Pa*ATPPRT:PRPP:histidine on the other hand have rmsd of 1.71, 1.37 and 1.75 respectively for the hetero-octamer, *Pa*HisZ tetramer and *Pa*HisG$_S$ dimers. Clearly, *Pa*ATPPRT:PRPP:histidine is not in the activated form and is similar in

structure to the non-activated conformation (when ATP is not bound). Perhaps, histidine locks *Pa*ATPPRT in this non-activated form, preventing conformational changes that would subsequently occur upon ATP binding.

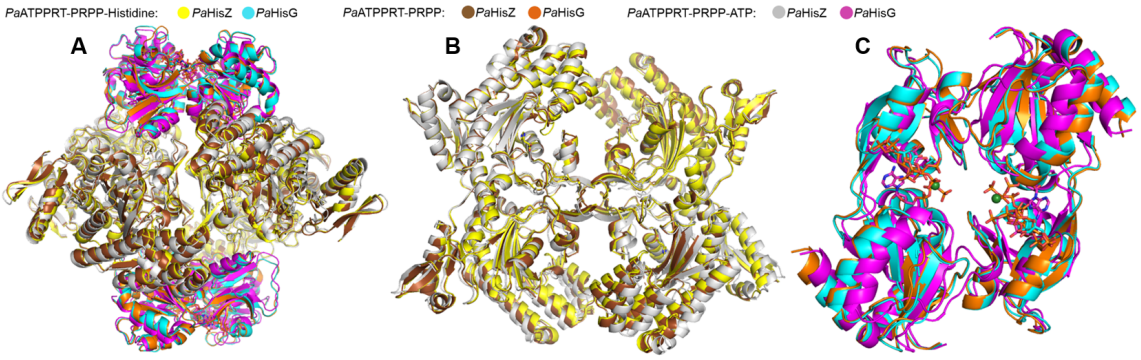


Figure 3.43: Overlay of *Pa*ATPPRT:PRPP, *Pa*ATPPRT:PRPP:histidine, and *Pa*ATPPRT:PRPP:ATP structures. (A) full hetero-octamers, (B) *Pa*HisZ tetramers and (C) *Pa*HisG$_S$ dimers.

When the *Pa*HisZ subunits from all published structures: *Pa*ATPPRT [20], *Pa*ATPPRT:PRPP, *Pa*ATPPRT:PRPP-ATP, *Pa*ATPPRT:PRATP [101] and *Pa*ATPPRT:PRPP:histidine Cα atoms are overlaid (Figure 3.44) two loops are seen to move substantially (shown by black arrows in Figure 3.44A) whereas the rest of the *Pa*HisZ structures vary very little from one the another. These loops (labelled as in the *Pa*ATPPRT:PRPP:histidine complex), span from Asp256 – Ile269 (the histidine binding loop) and from Asp101 - Leu117 at the *Pa*HisZ/*Pa*HisG$_S$ interface. The latter loop (Asp101 – Leu117) differs by over 2 Å in the *Pa*ATPPRT:PRPP:ATP activated conformation (Figure 3.44, grey) compared to all the other published *Pa*ATPPRT structures (Figure 3.44B, right), which are non-activated. The former loop (Asp256-Ile269) differs by over 4.5 Å between the *Pa*ATPPRT:PRPP:histidine structure (Figure 3.44, yellow) and all other published structures (Figure 3.44B, left). This includes an alternative His266 rotamer (for interaction with histidine (Figure 3.42) and a 4 Å shift of Tyr265 to form a hydrogen bond with the inhibitor. This results in the Tyr265 4-OH group residing 6 Å away (Figure 3.44B, yellow) from its position in all other structures.

In the *Pa*ATPPRT:PRPP:histidine structure, these two loops are connected by a hydrogen bond between Tyr263 (histidine binding loop) and His104 (Asp101-Leu117 at the *Pa*HisZ/*Pa*HisG$_S$ interface). This connection, which requires the tyrosine to rotate almost 180° from its most common position, is not seen in any of the other *Pa*ATPPRT structures.

When *Pa*ATPPRT:PRPP:histidine (an unactivated form) and *Pa*ATPPRT:PRPP:ATP (activated form) are overlaid (Figure 3.45) the activated form's Asp101-Leu117 interface

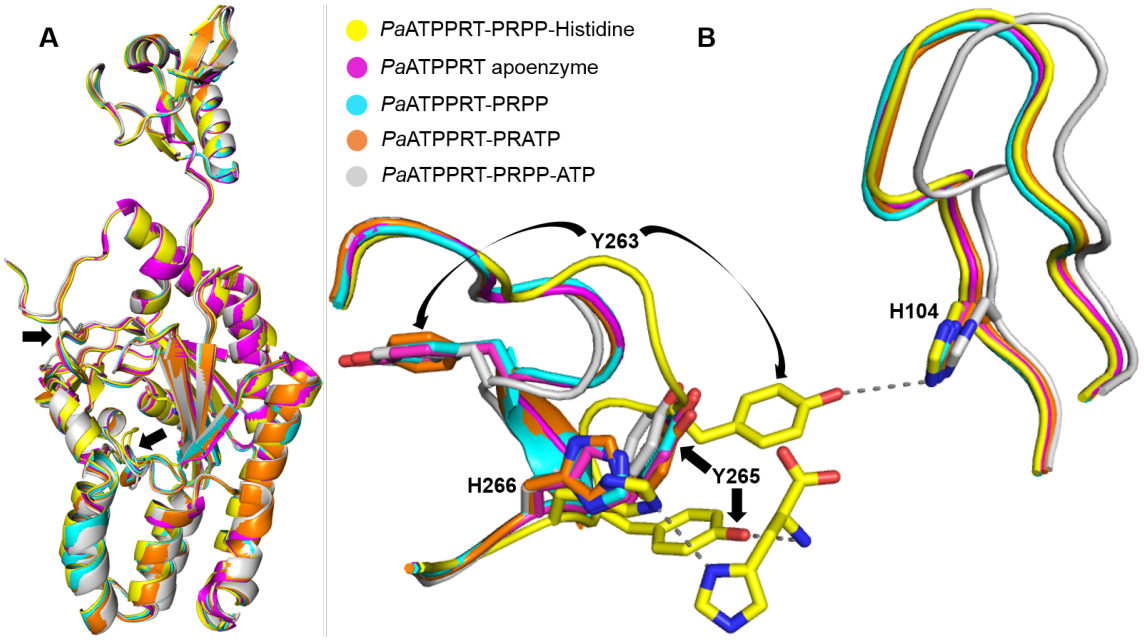


Figure 3.44: Overlay of *Pa*HisZ subunit conformations of all *Pa*ATPPRT structures. (A) Overlay of *Pa*HisZ subunits from all *Pa*ATPPRT structures (black arrows show the two loops with substantial movement) and (B) enlarged view of the two loops subject to large spatial movement. *Pa*HisZ residue side chains and bound histidine are shown in stick models and dashed lines show polar interactions. Black arrows highlight the different positions adopted by the same residue side chains in the different structures.

loop has residues Leu110 and Phe111 ideally positioned for hydrophobic interaction with the *Pa*HisG$_S$ monomer. Whereas in the inhibited structure, this loop is positioned towards the other *Pa*HisG$_S$ monomer which situates Leu110 in a position capable of disrupting its Glu82-His103 interaction. In addition, the inhibited structure also sees the aromatic ring of Phe111 shift closer to the interface between the two *Pa*HisG$_S$ subunits further disrupting the hydrophobic interactions made here, in the activated complex.

These interactions likely prevent the efficient sampling by the Asp101–Leu117 loop of the conformation that triggers catalytic activation when ATP binds to the *Pa*ATPPRT:PRPP complex in the absence of histidine. Perhaps it is also these changes in the interactions between the *Pa*HisZ loop and *Pa*HisG$_S$ monomers that are responsible, at least in part, for transmitting the allosteric signal from the regulatory subunit (*Pa*HisZ) to the active site in the catalytic subunit (*Pa*HisG$_S$). Nevertheless, the binding of inhibitor in *Pa*HisZ directly causes a shift of the α-helix and β-strand which connects the loop responsible for cross-subunit stabilisation of leaving group departure near the active site in *Pa*HisG$_S$, further from where the adjacent catalytic subunit resides in the activated form (Figure 3.45, black arrows). Taken together, these differences between the activated form of *Pa*ATPPRT:PRPP:ATP and the *Pa*ATPPRT:PRPP:histidine structures clearly demonstrate structural changes that are directly due to allosteric binding. Consequently, in the case of *Pa*ATPPRT, allosteric binding reduces the likelihood of sampling the

activated conformation upon ATP binding, thus stochastically restricting the structure to a non-activated form.

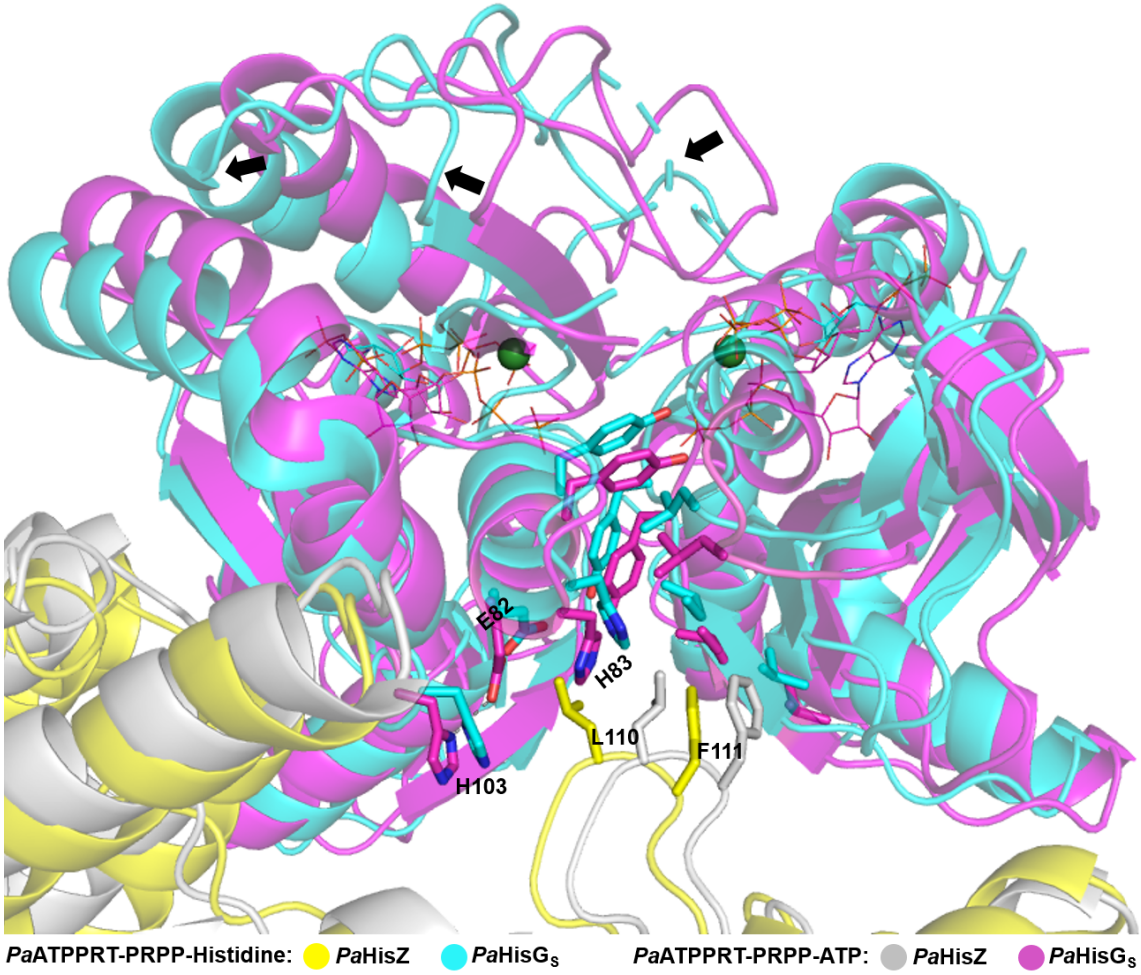


Figure 3.45: Ribbon diagram of overlaid *Pa*ATPPRT:PRPP:histidine and *Pa*ATPPRT:PRPP:ATP structures detailing part of one *Pa*HisZ and one of the *Pa*HisG$_S$ dimers at the *Pa*HisZ-*Pa*HisG$_S$ and *Pa*HisG$_S$-*Pa*HisG$_S$ interfaces. The figure shows substrates as wire frame, histidine and relevant side chains in stick models (carbon corresponding to subunit ribbon colour), nitrogen atoms (blue), oxygen atoms (red), protein backbone (ribbon diagram). Black arrows denote the distinct positions of the *Pa*HisG$_S$ loop responsible for leaving group stabilisation at the transition state.

Chapter 4

CONCLUDING DISCUSSION AND FURTHER WORK

This work demonstrates the first example of histidine binding in a site analogous to that of HisRS in a short-form ATPPRT. The resulting interaction between two loops in *Pa*HisZ mediated by a hydrogen bond between Tyr263 and His104 is likely to greatly reduce the probability of the complex sampling the activated conformation necessary for successful reaction despite the ability of histidine to bind to the free enzyme, the *Pa*ATPPRT:PRPP bound state and the *Pa*ATPPRT:PRPP:ATP bound state. Additionally, this work highlights a clear distinction between the activated form with ATP bound and all other structures. Although the ATP bound structure previously published shows ATP bound in a non-catalytically competent conformation (which could influence enzyme dynamics) the majority, if not all, of the interactions established are likely to still be present in the catalytic rotamer. Overall, this work demonstrates a clear reliance on structural changes for the allosteric regulation of *Pa*ATPPRT.

Although clear, the changes observed are not dramatic and this may be due to the restriction of movement imposed by crystal packing and the potentially limiting nature of exposing the crystal to histidine by soaking. It may be that the effects detailed here are more pronounced in solution but selected key mutations of Tyr265, Tyr263, Leu110 and Phe111 will highlight the importance of the interactions theorised here. Of these mutations, Tyr265Phe is currently underway within the da Silva group to determine whether removal of this interaction interrupts transmission of the allosteric signal and renders the enzyme insensitive to histidine inhibition. If successful, this gives scope for biotechnological advancement and drug discovery. Since *P. arcticus* and *A. baumannii* both contain a short-form ATPPRT (with 43% sequence alignment for HisZ and 69% for HisG), confirming the importance of the residues involved in binding and allosteric signal relay in *Pa*ATPPRT may provide insight into ways to design allosteric modulators

of *A. baumannii* ATPPRT for antibiotic development. Furthermore, creating a mutant of *Pa*ATPPRT capable of histidine production insensitive to feedback inhibition but still profiting from the activation of *Pa*HisG$_S$ by *Pa*HisZ has clear advantages for industrial histidine production. In addition, this work has confirmed ADP as a viable alternative substrate for *Pa*ATPPRT, with no indication of substrate inhibition, which complements this biosynthesis strategy.

Overall, this work provides the most detailed account of the allosteric inhibition mechanism of a short-form ATPPRT.

Bibliography

1. Monod, J. *Chance and Necessity: Essay on the Natural Philosophy of Modern Biology* (Vintage Books, New York, 1972).
2. Fenton, A. W. Allostery: an illustrated definition for the 'second secret of life'. *Trends in Biochemical Sciences* **33,** 420–5 (2008).
3. Makhlynets, O. V., Raymond, E. A. & Korendovych, I. V. Design of Allosterically Regulated Protein Catalysts. *Biochemistry* **54,** 1444–1456 (2015).
4. Grover, A. K. Use of Allosteric Targets in the Discovery of Safer Drugs. *Medical Principles and Practice* **22,** 418–426 (2013).
5. Reece, J. B. & Campbell, N. A. *Campbell Biology* 11th ed. (Pearson, Boston, 2011).
6. McMurry, J. & Begley, T. *The Organic Chemistry of Biological Pathways* 295–296 (Roberts and Company Publishers, Colarado 80111 USA, 2005).
7. Alfonso, M., Okondo, M. & Goodey, N. M. Examining how allosteric mutations affect ligand binding and specificity on dihydrofolate reductase. *Biochemistry and Molecular Biology* **31,** e760.10 (2017).
8. Goodey, N. M. & Benkovic, S. J. Allosteric regulation and catalysis emerge via a common route. *Nature Chemical Biology* **4,** 474–482 (2008).
9. Grispo, M. T., Natarajan, C., Projecto-Garcia, J., Moriyama, H., Weber, R. E. & Storz, J. F. Gene duplication and the evolution of hemoglobin isoform differentiation in birds. *The Journal of Biological Chemistry* **287,** 37647–58 (2012).
10. Meglasson, M. D., Schinco, M. & Matschinsky, F. M. Mannose Phosphorylation by Glucokinase from Liver and Transplantable Insulinoma: Cooperativity and Discrimination of Anomers. *Diabetes* **32,** 1146–1151 (1983).
11. Notides, A. C., Lerner, N & Hamilton, D. E. Positive cooperativity of the estrogen receptor. *Proceedings of the National Academy of Sciences of the United States of America* **78,** 4926–30 (1981).
12. Levit, A. & Koshland, D. E. Negative Cooperativity in Regulatory Enzymes. *Proceedings of the National Academy of Sciences of the United States of America* **62,** 1121–1128 (1969).

13. Clark, D. E., Poteet-Smith, C. E., Smith, J. A. & Lannigan, D. A. Rsk2 allosterically activates estrogen receptor alpha by docking to the hormone-binding domain. *The EMBO Journal* **20,** 3484–3494 (2001).

14. Velez, N., Brautigam, C. A. & Phillips, M. A. *Trypanosoma brucei* S-adenosylmethionine decarboxylase N terminus is essential for allosteric activation by the regulatory subunit prozyme. *The Journal of Biological Chemistry* **288,** 5232–40 (2013).

15. Auzat, I., Le Bras, G. & Garel, J.-R. The cooperativity and allosteric inhibition of *Escherichia coli* phosphofructokinase depend on the interaction between threonine-125 and ATP (phosphoryl transfer/site-directed mutgenesis/enzyme regulation). *Proceedings of the National Academy of Sciences of the United States of America* **91,** 5242–5246 (1994).

16. Berg, J. M. J. M., Tymoczko, J. L. & Stryer, L. *Biochemistry* (W.H. Freeman, New York, 2002).

17. Gerhart, J. From feedback inhibition to allostery: the enduring example of aspartate transcarbamoylase. *FEBS* **281,** 612–620 (2014).

18. Yatest, R. A. & Pardee, A. B. Control of pyrimidine biosynthesis in *Escherichia coli* by a feed-back mechanism. *Journal of Biological Chemistry* **221,** 757–770 (1956).

19. Curien, G., Job, D., Douce, R. & Dumas, R. Allosteric Activation of Arabidopsis Threonine Synthase by S -Adenosylmethionine. *Biochemistry* **37,** 13212–13221 (1998).

20. 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., *et al.* Kinetics and Structure of a Cold-Adapted Hetero-Octameric ATP Phosphoribosyltransferase. *Biochemistry* **56,** 793–803 (2017).

21. Thomson, J. A. The role of biotechnology for agricultural sustainability in Africa. *Philosophical Transactions of the Royal Society B* **363,** 905–913 (2008).

22. Wenda, S., Illner, S., Mell, A. & Kragl, U. Industrial biotechnology-the future of green chemistry? *Green Chemistry* **13,** 3007–3047 (2011).

23. Becker, J. & Wittmann, C. Systems and synthetic metabolic engineering for amino acid production - the heartbeat of industrial strain development. *Current Opinion in Biotechnology* **23,** 718–726 (2012).

24. Wendisch, V. F. Microbial production of amino acids and derived chemicals: Synthetic biology approaches to strain development. *Current Opinion in Biotechnology* **30,** 51–58 (2014).

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

26. Umbarger, H. E. Evidence for a negative-feedback mechanism in the biosynthesis of isoleucine. *Science* **123,** 848 (1956).

27. Kumar Grover, A. Use of Allosteric Targets in the Discovery of Safer Drugs. *Medical Principles and Practice* **22,** 418–426 (2013).

28. Wenthur, C. J., Gentry, P. R., Mathews, T. P. & Lindsley, C. W. Drugs for Allosteric Sites on Receptors. *Annu Rev Pharmacol Toxicol.* **54,** 165–184 (2014).

29. *Putative Cognitive Enhancer VU319* https://clinicaltrials.gov/ct2/show/NCT03220295 (2019).

30. *Phase I Dose Escalation Study With an Allosteric AKT 1/2 Inhibitor in Patients* https://clinicaltrials.gov/ct2/show/NCT01915576 (2019).

31. *Web of Science search for 'allosteric drug'* https://apps.webofknowledge.com (2017).

32. Abdel-Magid, A. F. Allosteric Modulators: An Emerging Concept in Drug Discovery. *ACS Med. Chem. Lett* **6,** 104–107 (2015).

33. Tee, W.-V., Guarnera, E. & Berezovsky, I. N. Reversing allosteric communication: From detecting allosteric sites to inducing and tuning targeted allosteric response. *PLOS Computational Biology* **14,** e1006228 (2018).

34. Lopez-munoz, F. & Uchaudabe, R. The history of barbiturates a century after their clinical introduction. *Neuropsychiatric Disease and Treatment* **1,** 329–343 (2005).

35. Löscher, W. & Rogawski, M. A. How theories evolved concerning the mechanism of action of barbiturates. *Epilepsia* **53,** 12–25 (2012).

36. Olsen, R. W. & DeLorey, T. M. *GABA Receptor Physiology and Pharmacology* 6th ed. (Lippincott-Raven, Philadelphia, 1999).

37. Pardee, A. B. & Reddy, G.-V. Beginnings of feedback inhibition, allostery, and multi-protein complexes. *Gene* **321,** 17–23 (2003).

38. Cui, Q. & Karplus, M. Allostery and cooperativity revisited. *Protein Science* **17,** 1295–1307 (2008).

39. Nussino, R. Introduction to Protein Ensembles and Allostery. *Chemical Reviews* **116,** 6263–6266 (2016).

40. Pauling, L. The Oxygen Equilibrium of Hemoglobin and Its Structural Interpretation. *Proceedings of the National Academy of Sciences of the United States of America* **21,** 186–91 (1935).

41. Koshland Jr, D., Nemethy, G & Filmer, D. Comparison of Experimental Binding Data and Theoretical Models in Proteins Containing Subunits. *Biochemistry* **5,** 365–385 (1966).

42. Vasconcelos, I. B., Basso, L. A. & Santos, D. S. Kinetic and Equilibrium Mechanisms of Substrate Binding to *Mycobacterium tuberculosis* Enoyl Reductase: Implications to Function-Based Antitubercular Agent Design. *J. Braz. Chem. Soc* **21,** 1503–1508 (2010).

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

44. Changeux, J.-P. & Edelstein, S. J. Allosteric mechanisms of signal transduction. *Science* **308,** 1424–8 (2005).

45. Kendrew, D. J. C., Bodo, G, Dintzis, H. M., Parrish, R. G., Wyckoff, H & Phillips, D. C. A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature* **181,** 662–666 (1958).

46. Cooper, A & Dryden, D. T. F. Allostery without conformational change. *European Biophysics Journal* **11,** 103–109 (1984).

47. Nussinov, R. & Tsai, C.-J. Allostery without a conformational change? Revisiting the paradigm. *Current Opinion in Structural Biology* **30,** 17–24 (2015).

48. Koshland Jr, D. E. Enzyme flexibility and enzyme action. *Journal of Cellular and Comparative Physiology* **54,** 245–258 (1959).

49. Whitley, M. J. & Lee, A. L. Frameworks for understanding long-range intra-protein communication. *Current protein & peptide science* **10,** 116–27 (2009).

50. Motlagh, H. N., Wrabl, J. O., Li, J. & Hilser, V. J. The ensemble nature of allostery. *Nature* **508,** 331–339 (2014).

51. Bryngelson, J. D. & Wolynes, P. G. Spin glasses and the statistical mechanics of protein folding (disordered systems/irreversible denaturation/molten-globule state/biomolecular self-assembly). *Proc. Nati. Acad. Sci. USA* **84,** 7524–7528 (1987).

52. Fisher, C. K. & Stultz, C. M. Constructing ensembles for intrinsically disordered proteins. *Current Opinion in Structural Biology* **21,** 426–431 (2011).

53. Agafonov, R. V., Wilson, C., Otten, R., Buosi, V. & Kern, D. Energetic dissection of Gleevec 's selectivity toward human tyrosine kinases. *Nature* **21,** 848–853 (2014).

54. Bloudek, L. M., Makenbaeva, D. & Eaddy, M. Anticipated Impact of Generic Imatinib Market Entry on the Costs of Tyrosine Kinase Inhibitors. *Am Health Drug Benefits.* **8,** 472–480 (2015).

55. Seeliger, M. A., Nagar, B., Frank, F., Cao, X., Henderson, M. N. & Kuriyan, J. c-Src Binds to the Cancer Drug Imatinib with an Inactive Abl/c-Kit Conformation and a Distributed Thermodynamic Penalty. *Structure* **15,** 299–311 (2007).

56. And, R. C. & Loria*, J. P. Evidence for Flexibility in the Function of Ribonuclease A. *Biochemistry* **41,** 6072–6081 (2002).

57. Nechushtai, R., Lammert, H., Michaeli, D., Eisenberg-Domovich, Y., Zuris, J. A., Luca, M. A., Capraro, D. T., Fish, A., Shimshon, O., Roy, M., *et al.* Allostery in the ferredoxin protein motif does not involve a conformational switch. *Proceedings of the National Academy of Sciences of the United States of America* **108,** 2240–5 (2011).

58. Pedreño, S., Pisco, J. P., Larrouy-Maumus, G., Kelly, G. & De Carvalho, L. P. S. Mechanism of feedback allosteric inhibition of ATP phosphoribosyltransferase. *Biochemistry* **51,** 8027–8038 (2012).

59. Carson, M., Johnson, D. H., Mcdonald, H., Brouillette, C. & Delucas, L. J. His-tag impact on structure. *Acta Crystallographica Section D* **63,** 295–301 (2007).

60. Gerhart, J. C. & Pardee, A. B. The Enzymology of Control by Feedback Inhibition. *The Journal of Biological Chemistry* **237,** 891–896 (1962).

61. Brooke, M. S., Ushiba, D. & Magasanik, B. Some Factors Affecting the Excretion of Orotic Acid By Mutants of *Aerobacter aerogenes*. *J. Bacteriol.* **68,** 534–540 (1954).

62. Novick, A. & Szilard, L. Experiments with the Chemostat on Spontaneous Mutations of Bacteria. *PNAS* **36,** 708–719 (1950).

63. Roberts, R. B., Abelson, P. H., Cowie, D. B., Bolton, E. T. & Britten, R. J. Studies of Biosynthesis in *Escherichia coli*. *The Quarterly Review of Biology* **31,** 155–156 (1956).

64. Yatest, R. A. & Pardee, A. B. Pyrimidine Biosynthesis in *Escherichia coli*. *Journal of Biological Chemistry* **221,** 743–756 (1956).

65. Banerjee, A., Wu, Y., Banerjee, R., Li, Y., Yan, H. & Sharkey, T. D. Feedback inhibition of deoxy-D-xylulose-5-phosphate synthase regulates the methylerythritol 4-phosphate pathway. *The Journal of Biological Chemistry* **288,** 16926–36 (2013).

66. Sugimoto, S.-i. & Shiio, I. Agricultural and Biological Chemistry Regulation of Tryptophan Biosynthesis by Feedback Inhibition of the Second-Step Enzyme, Anthranilate Phosphoribosyl-transferase, in *Brevibacterium flavum*. *Agricultural and Biological Chemistry* **47,** 2295–2305 (1983).

67. Twarog, R. Enzymes of the isoleucine-valine pathway in *Acinetobacter*. *Journal of bacteriology* **111,** 37–46 (1972).

68. Martin, R. G. The First Enzyme in Histidine Biosynthesis: The Nature of Feedback Inhibition by Histidine. *The Journal of Biological Chemistry* **238,** 257–268 (1963).

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

70. Champagne, K. S., Sissler, M., Larrabee, Y., Doublié, S. & Francklyn, C. S. Activation of the hetero-octameric ATP phosphoribosyl transferase through subunit interface rearrangement by a tRNA synthetase paralog. *The Journal of Biological Chemistry* **280,** 34096–34104 (2005).

71. Winkler, M. E. & Ramos-Montañez, S. Biosynthesis of Histidine. *EcoSal Plus* **3** (2009).

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

73. Brenner, M. & Ames, B. N. *The Histidine Operon and Its Regulation* 3rd ed., 349–387 (Elsevier Academic Press, Los Angeles, 1971).

74. Smith, J. L. Forming and Inhibiting PRT Active Sites. *Nat. Struc. Biol.* **6,** 502–504 (1999).

75. Morton, D. P. & Parsons, S. M. Biosynthetic direction substrate kinetics and product inhibition studies on the first enzyme of histidine biosynthesis, adenosine triphosphate phosphoribosyltransferase. *Arch. Biochem. Biophys.* **175,** 677–686 (1976).

76. Morton, D. P. & Parsons, S. M. Synergistic inhibition of ATP phosphoribosyltransferase by guanosine tetraphosphate and histidine. *Biochemical and Biophysical Research Communications* **74,** 172–177 (1977).

77. Pease, A. J., Roa, B. R., Luo, W. & Winkler, M. E. Positive growth rate-dependent regulation of the pdxA, ksgA, and pdxB genes of Escherichia coli K-12. *Journal of bacteriology* **184,** 1359–69 (2002).

78. Alifano, P., Fani, R., Lio', P., Lio', L., Lazcano, A., Bazzicalupo, M., Carlomagno, M. S. & Bruni, C. B. Histidine Biosynthetic Pathway and Genes: Structure, Regulation, and Evolution. *Microbiological reviews* **60,** 44–69 (1996).

79. Johnston, H. M., Barnest, W. M., Chumley, F. G., Bossi, L. & Roth, J. R. Model for regulation of the histidine operon of Salmonella. *Genetics* **77,** 508–512 (1980).

80. Malykh, E. A., Butov, I. A., Ravcheeva, A. B., Krylov, A. A., Mashko, S. V. & Stoynova, N. V. 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,** 42 (2018).

81. Mizukami, T., Hamu, A., Ikeda, M., Oka, T. & Katsumata, R. Cloning of the ATP Phosphoribosyl Transferase Gene of *Corynebacterium glutamicum* and Application of the Gene to L-Histidine Production. *Bioscience, Biotechnology, and Biochemistry* **58,** 635–638 (1994).

82. Zhang, Y., Shang, X., Deng, A., Chai, X., Lai, S., Zhang, G. & Wen, T. Genetic and biochemical characterization of Corynebacterium glutamicum ATP phosphoribosyltransferase and its three mutants resistant to feedback inhibition by histidine. *Biochimie* **94,** 829–838 (2012).

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

84. Vega, M. C., Zou, P., Fernandez, F. J., Murphy, G. E., Sterner, R., Popov, A. & Wilmanns, M. Regulation of the hetero-octameric ATP phosphoribosyl transferase complex from *Thermotoga maritima* by a tRNA synthase-like subunit. *Molecular Microbiology* **55,** 675–686 (2005).

85. Sissler, M, Delorme, C, Bond, J, Ehrlich, S. D., Renault, P & Francklyn, C. An aminoacyl-tRNA synthetase paralog with a catalytic role in histidine biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America* **96,** 8985–90 (1999).

86. Mittelstädt, G., Moggré, G.-J., Panjikar, S., Nazmi, A. R. & Parker, E. J. *Campylobacter jejuni* adenosine triphosphate phosphoribosyltransferase is an active hexamer that is allosterically controlled by the twisting of a regulatory tail. *Protein Science* **25,** 1492–506 (2016).

87. Bovee, M. L., Champagne, K. S., Demeler, B. & Francklyn, C. S. The Quaternary Structure of the HisZ - HisG N -1- (5 -Phosphoribosyl) -ATP Transferase from *Lactococcus lactis*. *Biochemistry* **41,** 11838–11846 (2002).

88. Song, J. & Markley, J. L. Cautionary tail: The presence of an N-terminal tag on dynein light-chain roadblock/lc7 affects its interaction with a functional partner. *Protein Pept. Lett.* **14,** 265–268 (2007).

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

90. Brashear, W. T. & Parsons, S. M. Evidence against a covalent intermediate in the adenosine triphosphate phosphoribosyltransferase reaction of histidine biosynthesis. *The Journal of Biological Chemistry* **250,** 6885–90 (1975).

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

92. 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. Uncoupling conformational states from activity in an allosteric enzyme. *Nature Communications* **8,** 1–9 (2017).

93. Cho, Y., Ioerger, T. R. & Sacchettini, J. C. Discovery of Novel Nitrobenzothiazole Inhibitors for *Mycobacterium tuberculosis* ATP Phosphoribosyl Transferase (HisG) through Virtual Screening. *Journal of Medicinal Chemistry* **51,** 5984–5992 (2008).

94. Pacholarz, K. J., Burnley, R. J., Jowitt, T. A., Ordsmith, V., Pisco, J. P., Porrini, M., Larrouy-Maumus, G., Garlish, R. A., Taylor, R. J., de Carvalho, L. P. S., *et al.* Hybrid Mass Spectrometry Approaches to Determine How L-Histidine Feedback Regulates the Enzyzme MtATP-Phosphoribosyltransferase. *Structure* **25,** 730–738.e4 (2017).

95. Quiocho, F. A. Atomic basis of the exquisite specificity of phosphate and sulfate transport receptors. *Kidney International* **49,** 943–946 (1996).

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

97. Ayala-del Rio, H. L., Chain, P. S., Grzymski, J. J., Ponder, M. A., Ivanova, N., Bergholz, P. W., Di Bartolo, G., Hauser, L., Land, M., Bakermans, C., *et al.* The Genome Sequence of *Psychrobacter arcticus* 273-4, a Psychroactive Siberian Permafrost Bacterium, Reveals Mechanisms for Adaptation to Low-Temperature Growth. *Applied and Environmental Microbiology* **76,** 2304–2312 (2010).

98. Ponder, M. A., Thomashow, M. F. & Tiedje, J. M. Metabolic activity of Siberian permafrost isolates, *Psychrobacter arcticus* and *Exiguobacterium sibiricum*, at low water activities. *Extremophiles* **12,** 481–490 (2008).

99. Ponder, M. A., Gilmour, S. J., Bergholz, P. W., Mindock, C. A., Hollingsworth, R., Thomashow, M. F. & Tiedje, J. M. Characterization of potential stress responses in ancient Siberian permafrost psychroactive bacteria. *FEMS Microbiology Ecology* **53,** 103–115 (2005).

100. Violot, S., Aghajari, N., Czjzek, M., Feller, G., Sonan, G. K., Gouet, P., Gerday, C., Haser, R. & Receveur-Bréchot, V. Structure of a Full Length Psychrophilic Cellulase from *Pseudoalteromonas haloplanktis* revealed by X-ray Diffraction and Small Angle X-ray Scattering. *Journal of Molecular Biology* **348,** 1211–1224 (2005).

101. Alphey, M. S., Fisher, G., Ge, Y., Gould, E. R., Machado, T. G., Liu, H., Florence, G. J., Naismith, J. H. & Guimaraes, R. Catalytic and anticatalytic snapshots of a short-form ATP phosphoribosyltransferase. *ACS Catalysis* **8,** 5601–5610 (2018).

102. Fisher, G., Thomson, C. M., Stroek, R., Czekster, C. M., Hirschi, J. S. & Silva, R. G. Allosteric Activation Shifts the Rate-Limiting Step in a Short-Form ATP Phosphoribosyltransferase. *Biochemistry* **57,** 4357–4367 (2018).

103. Brashear, W. T. & Parsons, S. M. Evidence against a covalent intermediate in the adenosine triphosphate phosphoribosyltransferase reaction of histidine biosynthesis. *The Journal of Biological Chemistry* **250,** 6885–90 (1975).

104. Cho, Y., Sharma, V. & Sacchettini, J. C. Crystal Structure of ATP phosphoribosyltransferase from *Mycobacterium tuberculosis*. *J Biol Chem* **278,** 8333–8339 (2003).

105. Sissler, M, Delorme, C, Bond, J, Ehrlich, S. D., Renault, P & Francklyn, C. An aminoacyl-tRNA synthetase paralog with a catalytic role in histidine biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America* **96,** 8985–90 (1999).

106. Moggré, G.-J., Poulin, M. B., Tyler, P. C., Schramm, V. L. & Parker, E. J. Transition State Analysis of Adenosine Triphosphate Phosphoribosyltransferase. *ACS Chemical Biology* **12,** 2662–2670 (2017).

107. Champagne, K. S., Sissler, M., Larrabee, Y., Doublié, S. & Francklyn, C. S. Activation of the hetero-octameric ATP phosphoribosyl transferase through subunit interface rearrangement by a tRNA synthetase paralog. *The Journal of Biological Chemistry* **280,** 34096–34104 (2005).

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

109. Morton, D. P. & Parsons, S. M. Inhibition of ATP phosphoribosyltransferase by AMP and ADP in the absence and presence of histidine. *Archives of Biochemistry and Biophysics* **181,** 643–648 (1977).

110. Lohkamp, B., McDermott, G., Campbell, S. A., Coggins, J. R. & Lapthorn, A. J. The Structure of *Escherichia coli* ATP-phosphoribosyltransferase: Identification of Substrate Binding Sites and Mode of AMP Inhibition. *Journal of Molecular Biology* **336,** 131–144 (2004).

111. Wang, N., Ozer, E. A., Mandel, M. J. & Hauser, R. Genome-Wide Identification of *Acinetobacter baumannii* Genes Necessary for Persistence in the Lung. **5,** 1–8 (2014).

112. Miroux, B. & Walker, J. E. Over-production of Proteins in *Escherichia coli*: Mutant Hosts that Allow Synthesis of some Membrane Proteins and Globular Proteins at High Levels. *J. Mol. Biol* **260,** 289–298 (1996).

113. Divan, A. & Royds, J. *Tools and techniques in biomolecular science* 156 (Oxford University Press, Oxford, 2013).

114. Agilent Technologies. *BL21Gold Competent Cells, BL21-Gold(DE3) Competent Cells, and BL21Gold(DE3)pLysS Competent Cells* `http:\\www.genomics.agilent.com` (2019).

115. Niesen, F. H., Berglund, H. & Vedadi, M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nature Protocols* **2,** 2212–2221 (2007).

116. Morrison, J. Kinetics of the reversible inhibition of enzyme-catalysed reactions by tight-binding inhibitors. *Biochimica et Biophysica Acta (BBA) - Enzymology* **185,** 269–286 (1969).

117. Reed, M. C., Lieb, A. & Nijhout, H. F. The biological significance of substrate inhibition: A mechanism with diverse functions. *BioEssays* **32,** 422–429 (2010).

118. Kabsch, W. XDS. *Acta crystallographica. Section D, Biological crystallography* **66,** 125–32 (2010).

119. Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta crystallographica. Section D, Biological crystallography* **69,** 1204–14 (2013).

120. Vagin, A. & Teplyakov, A. MOLREP : an Automated Program for Molecular Replacement. *Journal of Applied Crystallography* **30,** 1022–1025 (1997).

121. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D Biological Crystallography* **60,** 2126–2132 (2004).

122. Murshudov, G. N., Vagin, A. A., Dodson, E. J. & IUCr. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallographica Section D Biological Crystallography* **53,** 240–255 (1997).

123. Magill, S. S., O'Leary, E., Janelle, S. J., Thompson, D. L., Dumyati, G., Nadle, J., Wilson, L. E., Kainer, M. A., Lynfield, R., Greissman, S., *et al.* Changes in Prevalence of Health Care–Associated Infections in U.S. Hospitals. *New England Journal of Medicine* **379,** 1732–1744 (2018).

124. Kourtis, A. P., Hatfield, K., Baggs, J., Mu, Y., See, I., Epson, E., Nadle, J., Kainer, M. A., Dumyati, G., Petit, S., *et al.* Vital Signs: Epidemiology and Recent Trends in Methicillin-Resistant and in Methicillin-Susceptible Staphylococcus aureus Bloodstream Infections — United States. *MMWR. Morbidity and Mortality Weekly Report* **68,** 214–219 (2019).

125. McGuinness, W. A., Malachowa, N. & DeLeo, F. R. Vancomycin Resistance in *Staphylococcus aureus*. *The Yale journal of biology and medicine* **90,** 269–281 (2017).

126. Henriksen, S., Liu, J., Estiu, G., Oltvai, Z. & Wiest, O. Identification of novel bacterial histidine biosynthesis inhibitors using docking, ensemble rescoring, and whole-cell assays. *Bioorganic & Medicinal Chemistry* **18,** 5148–5156 (2010).

127. Larentis, A. L., Nicolau, J. F. M. Q., Esteves, G. d. S., Vareschini, D. T., de Almeida, F. V. R., dos Reis, M. G., Galler, R. & Medeiros, M. A. Evaluation of pre-induction temperature, cell growth at induction and IPTG concentration on the expression of a leptospiral protein in *E. coli* using shaking flasks and microbioreactor. *BMC Research Notes* **7,** 671 (2014).

128. Genlantis. *SoluBL21™ Competent E. coli* `http://www.genlantis.com/commerce/ccp1779-1029-solubl21-competent-e--coli-sb21{\_}cp18.htm` (2019).

129. Yang, Q., Xu, J., Li, M., Lei, X. & An, L. High-level expression of a soluble snake venom enzyme, gloshedobin, in *E. coli* in the presence of metal ions. *Biotechnology Letters* **25,** 607–610 (2003).

130. Kapust, R. B. & Waugh, D. S. *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Science : A Publication of the Protein Society* **8,** 1668 (1999).

131. Seyfi Abad Shapouri, M. R., Mahmoodi, P., Gharibi, D., Ghorbanpoor, M., Yaghoubi, S., Rezaei, E., Rashno, M. & Mehravar, N. Molecular Cloning, Expression and Peroxidase Conjugation of *Staphylococcus aureus* Protein A. *Iranian journal of biotechnology* **14,** 230–235 (2016).

132. Nelp, M. T., Young, A. P., Stepanski, B. M. & Bandarian, V. Human Viperin Causes Radical SAM-Dependent Elongation of *Escherichia coli*, Hinting at Its Physiological Role. *Biochemistry* **56,** 3874–3876 (2017).

133. Ghssein, G., Brutesco, C., Ouerdane, L., Fojcik, C., Izaute, A., Wang, S., Hajjar, C., Lobinski, R., Lemaire, D., Richaud, P., *et al.* Biosynthesis of a broad-spectrum nicotianamine-like metallophore in *Staphylococcus aureus*. *Science* **352,** 1105–1109 (2016).

134. Lipsky, M. & Parsons, S. M. A Unique Reactive Residue in Adenosine Triphosphate Phosphoribosyltransferase Sensitive to Five Conformation and Dissociation States. *J Biol Chem* **250,** 5660–5668 (1975).

135. Kühl, P. W. Excess-substrate inhibition in enzymology and high-dose inhibition in pharmacology: a reinterpretation. *The Biochemical journal* **298,** 171–180 (1994).

136. Bhagavan, N. V. & Ha, C.-E. *Essentials of medical biochemistry : with clinical cases* 47–58 (Elsevier/Academic Press, 2011).

137. Nienhaus, K., Deng, P., Belyea, J., Franzen, S. & Nienhaus, G. U. Spectroscopic study of substrate binding to the carbonmonoxy form of dehaloperoxidase from *Amphitrite ornata*. *J. Phys.Chem. B* **110,** 13264–13276 (2006).

138. Yao, L., Vogeli, B., Torchia, D. A. & Bax, A. Simultaneous NMR study of protein structure and dynamics using conservative mutagenesis. *J. Phys. Chem. B* **112,** 6045–6056 (2008).

139. Freydank, A. C., Brandt, W. & Drager, B. Protein structure modeling indicates hexahistidine-tag interference with enzyme activity. *Proteins: struct.,Funct., Bioinf.* **72,** 173–183 (2008).

140. Chant, A., Kraemer-Pecore, C. M., Watkin, R. & Kneale, G. G. Attachment of a histidine tag to the minimal zinc nger protein of the *Aspergillus nidulans* gene regulatory protein area causes a conformational change at the DNA-binding site. *Protein Expression Purif.* **39,** 152–159 (2005).

141. Haldane, J. *Enzymes* (eds Plimmer, R. & Hopkins, F. G.) (Longmans, Green & Co., London, 1930).

142. Copeland, R. A. *Evaluation of Enzyme Inhibitors in Drug Discovery* 117–121 (Wiley, Hoboken, 2013).

143. Brenner, M. & Ames, B. N. *Metabolic Regulation* (ed Voge, H. J.) 349–387 (Academic Press, New York, 1971).

144. Lohkamp, B., McDermott, G., Campbell, S. A., Coggins, J. R. & Lapthorn, A. J. The structure of *Escherichia coli* ATP-phosphoribosyl transferase: identification of substrate binding sites and mode of AMP inhibition. *J Mol Biol* **336,** 131–144 (2004).

145. Duke University. *MolProbity* `http://molprobity.biochem.duke.edu/` (2019).