

The Search for Antibiotic Leads:  
Targeting Pantothenate Kinase in  
*Pseudomonas aeruginosa*



Samuel David Sutton

Department of Biomedicine  
University of Bergen

This thesis is submitted in partial fulfilment of the requirements for the  
degree of

*Master of Science in Biomedical Sciences*

Spring 2021



# Acknowledgements

I would like to thank my co-supervisors Charis Georgiou and Ruth Brenk who have provided excellent guidance throughout my Master's research. I am especially grateful for their willingness to accommodate my requests for clarification or assistance at odd hours and on their own free time. The theoretical and experimental knowledge they provided me greatly contributed to my understanding of and ability to complete my thesis. A previous Master's student, Ying Ho Leon Lee, conducted considerable work prior to the writing of this thesis without which my results could not have been made possible. I would also like to show appreciation for Ludvik Olai Espeland for his assistance in understanding the fragment screening process. My special thanks to Khanh Kim Dao, who often offered assistance with preparing or obtaining necessary experiment components, as well as the rest of the Brenk lab group for their unhesitating willingness to help and offer advice.





# Summary

Antibiotic resistance is a growing problem around the world, and poses difficulties in treating bacterial infections in a hospital setting. Routine operations are becoming riskier as the threat of hard to treat opportunistic infection rises. *Pseudomonas aeruginosa* has worryingly acquired carbapenem resistance, and such strains have been granted "critical" priority for new research by the World Health Organization (WHO), as its natural antibiotic resistance and newly observed carbapenem resistance means that there could soon be no effective treatments remaining. It is a ubiquitous Gram-negative bacterium now frequently found in hospitals. *P. aeruginosa* accounts for an estimated 10-15% of hospital-acquired infections globally and is particularly prevalent in clinical isolates from ventilator-associated pneumonia (VAP) and cystic fibrosis patients.

In this work, heavy focus was placed on the *P. aeruginosa* type III pantothenate kinase, a 27 kDa protein encoded for by the *coaX* gene which is involved in the first step of coenzyme A biosynthesis. Unlike many bacterial species, *P. aeruginosa* possesses only a single pantothenate kinase gene. The loss of PanK activity has been shown to prevent growth and reproduction in multiple bacterial species. Inhibitors of other PanK proteins have not been shown to inhibit growth in *P. aeruginosa*, likely owing to its considerable structural divergence from other characterized PanK isoforms. Despite this, a genome-wide ligandability screen using the previously published data of a selenium methionine containing *PaPanK* identified *PaPanK* as a potentially druggable target. This coupled with *PaPanK*'s necessity for *P. aeruginosa* growth makes it an attractive target for antibiotic development.

In order to eventually develop antibiotics against carbapenem resistant *P. aeruginosa* (CRPA), we have implemented a medium throughput fragment screening methodology,

as well as methods for the purification and crystallization of *PaPanK*. Fragment screening using Bio-layer interferometry (BLI) was conducted in order to probe the active site of *PaPanK* and resulted in identification of a number of compounds with micromolar dissociation constants ( $K_D$ ). Protein crystals, containing the *PaPanK* apoenzyme and complexes with a variety of ligands, were diffracted at X-ray beamline facilities. From these, the first crystal structures of *PaPanK* and *PaPanK*-Pantothenate with the wild type methionine rather than selenium methionine were determined. Additionally, diffraction data was obtained which indicated electron density for ATP binding, as well as a previous virtual screening hit.

# Abbreviations

---

ADP	Adenosine diphosphate
AMP-PNP	Adenylyl-imidodiphosphate
ATP	Adenosine triphosphate
Avi	Avidin tag
BLI	Bio-layer interferometry
BirA	Biotin ligase
CM	Cryomix
CRPA	Carbapenem resistant <i>Pseudomonas aeruginosa</i>
CV	Column volume
CoA	Coenzyme A
DMSO	Dimethyl sulfoxide
DTT	Dithiothreitol
Da	Dalton
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylenediaminetetraacetic acid
EU-OS	EU-OPENSREEN
GST	Glutathione-s-Transferase
IMAC	Immobilized metal affinity chromatography
IPTG	Isopropyl $\beta$ - d-1-thiogalactopyranoside
kDa	Kilodalton
LB	Lysogeny broth
MW	Molecular weight
PAGE	Polyacrylamide gel electrophoresis
PDB	Protein database
PEG	Polyethylene glycol
<i>PaPanK</i>	<i>Pseudomonas aeruginosa</i> pantothenate kinase
RCF	Relative centrifugal force
RPM	Revolutions per minute
Rmsd	Root mean square deviation
SDS	Sodium dodecyl sulfate
SEC	Size exclusion chromatography
SPR	Surface plasmon resonance
SSA	Super Streptavidin (biosensor)
Se-Met	Selenomethionine
TEV –	Tobacco etch virus
<i>Tm</i>	<i>Thermotoga maritima</i>
Tris	Tris (hydroxymethyl) aminomethane
VAP	Ventilator-associated pneumonia
WHO	World Health Organization

# Table of contents

Acknowledgements	iii
Summary	v
Abbreviations	vii
<b>1 Introduction</b>	<b>1</b>
1.1 The need for new antibiotics . . . . .	1
1.2 <i>P. aeruginosa</i> infection . . . . .	1
1.3 Antibiotic-resistance mechanisms of <i>P. aeruginosa</i> . . . . .	2
1.4 Coenzyme A and the role of pantothenate kinase . . . . .	4
1.5 <i>PaPanK</i> as a target for antibiotics . . . . .	6
1.6 Fragment screening . . . . .	6
1.7 X-ray crystallography . . . . .	9
<b>2 Aims</b>	<b>11</b>
<b>3 Methods</b>	<b>13</b>

---

3.1	Buffer preparation . . . . .	13
3.2	SDS-PAGE . . . . .	13
3.3	Plasmids utilized . . . . .	14
3.4	Transformation of <i>E. coli</i> with 6HisTEV <i>PaPank</i> and 6HisAvi <i>PaPank</i> .	14
3.5	Glycerol stock preparation . . . . .	15
3.6	Lysis and purification of recombinant <i>PaPank</i> cell pellets . . . . .	15
3.7	6His-tag cleavage and reverse His-Trap purification . . . . .	16
3.8	GST-BirA expression and purification . . . . .	17
3.9	Enzymatic biotinylation of 6HisAvi <i>PaPank</i> . . . . .	18
3.10	Bio-layer interferometry . . . . .	18
3.10.1	Protein loading optimization . . . . .	19
3.10.2	Compound binding . . . . .	20
3.11	<i>PaPank</i> crystallization, ligand soaking, and structural determination .	21
3.11.1	Cryoprotectant selection . . . . .	22
3.11.2	Soaking of <i>PaPank</i> crystals with ligands . . . . .	23
3.11.3	Alternative soaking of <i>PaPank</i> crystals with ligands . . . . .	23
3.11.4	Structural determination and refinement of X-ray data . . . . .	23
3.12	Work contributed by others . . . . .	24
<b>4</b>	<b>Results</b>	<b>25</b>
4.1	Expression and Purification of 6HisAvi <i>PaPank</i> . . . . .	25
4.2	BLI with <i>PaPank</i> . . . . .	26

---

4.2.1	Protein loading optimization . . . . .	26
4.2.2	Establishment of ATP and pantothenate binding . . . . .	27
4.2.3	Testing of virtual screening hits . . . . .	29
4.2.4	Fragment screen of EU-OS library . . . . .	33
4.2.5	Fragment screen of carboxylic acid library . . . . .	37
4.3	Discussion . . . . .	41
4.3.1	Expression and purification of 6HisAvi <i>PaPanK</i> . . . . .	41
4.3.2	BLI with <i>PaPanK</i> . . . . .	42
4.3.3	Protein loading optimization . . . . .	42
<b>5</b>	<b><i>PaPanK</i> protein crystallography</b>	<b>47</b>
5.1	Expression and Purification of 6HisTEV <i>PaPanK</i> . . . . .	47
5.1.1	TEV protease cleavage of 6HisTEV <i>PaPanK</i> . . . . .	48
5.2	Crystallization of <i>PaPanK</i> . . . . .	49
5.2.1	Determination of suitable cryoprotectant . . . . .	51
5.2.2	Diffraction and structural confirmation of <i>PaPanK</i> and various ligands . . . . .	52
5.3	Discussion: Protein crystallography . . . . .	57
5.3.1	Expression and purification of 6HisTEV <i>PaPanK</i> . . . . .	57
5.3.2	Crystallization of <i>PaPanK</i> . . . . .	57
5.3.3	Diffraction and structural confirmation of <i>PaPanK</i> and various ligands . . . . .	58

<b>6 Conclusion</b>	<b>61</b>
<b>References</b>	<b>63</b>
<b>Appendix A</b>	<b>71</b>



# Chapter 1

## Introduction

### 1.1 The need for new antibiotics

While a wide array of antibiotics exist, the rate at which new classes of antibiotics are discovered has dropped dramatically. The vast majority of antibiotic classes currently in use were discovered prior to 1970 and many potential new classes are of limited effect against most Gram-negative bacteria, which means that the situation remains dire [1]. One such Gram-negative species is *Pseudomonas aeruginosa* (*P. aeruginosa*), the carbapenem-resistant strains of which have been deemed "critical" priority for research by the World Health Organization (WHO) and are the targets of this work [2].

### 1.2 *P. aeruginosa* infection

*P. aeruginosa* is an aerobic, rod-shaped Gram-negative bacterium from the  $\gamma$ -proteobacteria class, and is of growing clinical significance due to its ability to develop resistance to most known antibiotics [3]. Capable of growing rapidly at temperatures ranging from 25-42C, it is ubiquitous in nature and is capable of causing disease in both plants and animals. Associated with high mortality rates, it has become an increasingly common cause of nosocomial infections of the lungs, bloodstream, urinary-tract, and burn/surgical sites [4][5]. It has been identified as one of the primary opportunistic infections in those with cystic fibrosis, and is associated with increased morbidity

and mortality amongst patients with chronic obstructive pulmonary disease (COPD) and bronchiectasis [6]. Though rarely causing pneumonia outside of hospitals, *P. aeruginosa* is a common cause of pneumonia within hospitals and is the leading cause of ventilator-associated pneumonia (VAP) [7].

*P. aeruginosa* possesses a number of virulence factors which contribute to its ability to infect and harm its host. A number of studies have examined the mechanisms by which *P. aeruginosa* initially attaches to surfaces and thereafter forms biofilms [8][9][10]. Following attachment, loss of the mechanosensory protein PilY1 modulates the activation of virulence and could be contributory to chronic infections.[11]. *P. aeruginosa* expresses a number of proteins which can be directly injected into adjacent eukarotic cells through the type III secretion system including exotoxin A (ExoA), ExoS, ExoT, ExoU, and ExoY [12]. The most potent toxin in *P. aeruginosa* is the protein *Pseudomonas* exotoxin A [PE], a member of the mono-ADP-ribosyltransferase family which inhibits eukaryotic protein synthesis and leads to cell-death [13].

### 1.3 Antibiotic-resistance mechanisms of *P. aeruginosa*

*P. aeruginosa* shares an inducible *ampC* gene mostly found in the Enterobacteriaceae bacterial family, encoding for a periplasmic  $\beta$ -lactamase protein capable of binding and hydrolyzing many cephalosporins and most penicillins, amongst other antibiotics to limited extents [14]. The highly-effective carbapenem class of antibiotics first emerged in the 1970s, belonging to the Beta-lactams but possessing greater inhibition against many resistant strains of bacteria including *P. aeruginosa*. Birnbaum et. al (1985) attributes the ability of carbapenems to resist  $\beta$ -lactamase activity moreso than other  $\beta$ -lactams to its structure, which contains a trans oriented hydroxyethyl side chain, rather than the typical cis acylamino side chain [15]. This class of antibiotics has expanded to include imipinem, panipenem, meropenem, biapenem, ertapenem, doripenem, and tebipenem, each of which possesses distinct pharmacokinetic and pharmacodynamic properties [16]. Carbapenems have been a common treatment for *P. aeruginosa* because of its normal inability to prevent their entrance, degrade, or efflux them to concentrations below the minimum inhibitory concentration (MIC) [17].

However, the resistance to the carbapenem imipenem has been known to occur in *P. aeruginosa* since at least 1989, and strains resistant to all carbapenems known as carbapenem resistant *P. aeruginosa* (CRPA) have since emerged [18][19]. Owing to the already limited permeability the *P. aeruginosa* outer membrane, mutational loss of the porin OprD (D2) is sufficient to confer imipenem resistance.  $\beta$ -metalloproteases are rarely found, but capable of hydrolyzing carbapenems, and the extended-spectrum cephalosporinases are also capable of limiting the effectiveness of carbapenems. Efflux pumps, an innate factor of *P. aeruginosa*'s general antibiotic resistance, can be overexpressed and contribute to carbapenem-resistance even in the absence of any carbapenem-hydrolyzing enzymes [20].

The development of CRPA strains has contributed to a worrying prevalence of clinically severe infections. Amongst antibiotic-resistant bacterial infections, CRPA-attributed bacterial deaths in the European Union trail only methicilin-resistant *Staphylococcus aureus* and third-generation cephalosporin-resistant *Escherichia coli* [21]. Indeed, carbapenem-resistance itself has been tentatively-linked to increased mortality when comparing *P. aeruginosa* and CRPA-mediated bacteremia [22].

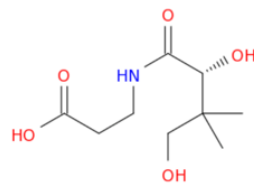
When carbapenem fails to treat CRPA infection, antibiotics with more potent side-effects such as polymyxin B and colistin (polymyxin E) are employed. Formerly considered too nephrotoxic for regular clinical use, carefully controlled doses of polymyxins have been used effectively against CRPA. The positively-charged polymyxins can mitigate CRPA's innate and acquired defenses by binding to the negatively-charged extracellular portion of Lipid A, passing through and acting in a detergent-like manner which fatally disrupts the outer membrane [23]. Yet still, resistance to even these 'last line of defense' treatments has been observed, with one study identifying colistin resistance in 21.3% of tested clinical isolates [24].

Recently, several clinical trials for antibiotics which could treat CRPA have taken place. With few exceptions, it is primarily members of existing antibiotic classes and combinatorial treatment with existing antibiotics that are in advanced stages of development [25]. Murepavidin of the novel outer membrane protein-targeting antibiotic (OMPTA) class has shown promising efficacy against CRPA, but resistance mechanisms have already been observed within clinical isolates [26]. Given the demonstrated ability of *P. aeruginosa* to develop resistance to the vast majority of existing antibiotics, it is of the utmost importance that this repertoire be expanded as much as possible.

## 1.4 Coenzyme A and the role of pantothenate kinase

CoenzymeA [CoA] is a cofactor found in and necessary for all known organisms to survive [27]. It consists of pantothenate linked to both a 3',5'-adenosine diphosphate and a cysteamine. The cysteamine moiety of CoA can be acetylated by acyl-CoA synthetases, producing the acetyl-CoA which serves as the acetyl carrier in the tricarboxylic acid cycle of aerobic organisms [28]. Acetyl-CoA is also a necessary component in fatty acid synthesis, where it is utilized directly by both fatty acid biosynthesis enzyme H (FabH) and the acetyl-CoA carboxylate complex (ACC) as initial steps of the synthesis pathway. Further modification of these fatty acids produces glycerolipids, key components of the cellular membranes found in both Gram-positive and negative bacteria [29]. Beyond that, CoA and compounds originating from CoA are estimated to be involved in at least 4% of all reactions within *E. coli* [30].

The biosynthesis of CoA is a multi-step process involving several enzymes, and despite often vast differences in amino acid sequence the overall pathway has remained conserved [31]. The first and rate-limiting step involves the phosphorylation of pantothenate (Figure 1.1) by pantothenate kinase (PanK), which produces the 4'-phosphopantothenate utilized later in the CoA synthesis pathway [27].

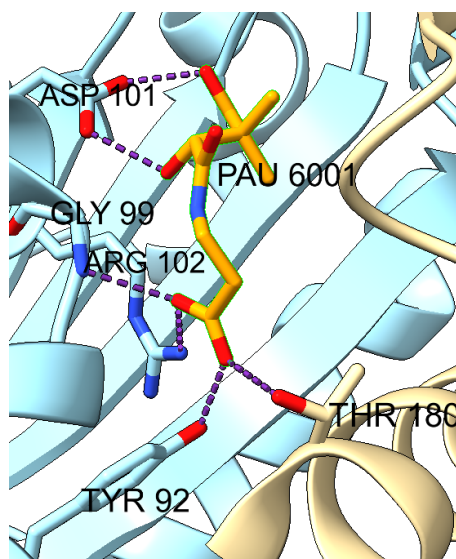


**Fig. 1.1:** Pantothenate, the substrate of PanKs including *PaPanK*, the phosphorylation of which is a vital step of CoA synthesis.

Three different types of bacterial PanKs are known to exist. In type I, the functional protein exists as a tetramer consisting of 4 identical subunits. Type I PanK is found only in Eubacteria and is encoded for by the *coaA* gene [32]. The type II and III PanKs exist as dimers, with type II being found in eukaryotes and a few bacterial species such as *Staphylococcus aureus*. All known type III PanKs and some type II PanKs are notable for their lack of feedback inhibition by CoA or its thioesters [33]. Type III PanKs represent the most recently classified isoforms of PanK and are encoded for

by the *coaX* gene. In contrast to type I and type II isoforms, type III PanKs have a closed substrate binding pocket and bind ATP much more weakly [34].

One such example of a type III PanK is the *P. aeruginosa* PanK (*PaPanK*), the sole protein exhibiting PanK activity in *P. aeruginosa*. It has been shown that in addition its typical requirement of  $Mg^{2+}$  for ATP binding, its open ATP binding pocket necessitates the presence of a singly charged cation for phosphorylation to occur. The activity of *PaPanK* is nearly zero when only  $Na^+$  or  $Li^+$  cations are used, with the presence of any of  $K^+$ ,  $Rb^+$ , or  $NH_4^+$  strongly preferred [35]. While *PaPanK*'s affinity for pantothenate ( $K_M=5 \mu M$ ) is similar to the type I and II PanKs, it has a much lower affinity for ATP ( $K_M=7.3 mM$ ). Given the extremely high concentration of ATP required for phosphorylation of pantothenate, Hong et al. (2006) tested the activity of *PaPanK* with a variety of alternative nucleotide phosphoryl donors and found that similarly high concentrations were required regardless of the phosphoryl donor used. In the same work, the structures of Se-Met *PaPanK* and Se-Met *PaPanK*-Pantothenate (Figure 1.2) were determined, with pantothenate seen binding at the dimer interface [35].



**Fig. 1.2:** The pantothenate-binding active site of Se-Met *PaPanK*-pantothenate reproduced from PDB 2F9W, showing the hydrogen bonds (dashed lines) between pantothenate (labelled PAU) and the surrounding protein residues. The two protein chains are colored differently.

## 1.5 *PaPanK* as a target for antibiotics

Antibiotics made to target specific species of pathogenic bacteria are highly appealing, as they can target necessary biological mechanisms to which there exists no known resistance. Additionally, such antibiotics could be less disruptive to the surrounding microbiome than traditional broad-spectrum antibiotics due to their species-specificity (note: need to access paper on campus before citing)[36]. Perhaps most importantly of all, species-specific antibiotics would avoid the risk of over-use that greatly contributes to the development of antibiotic resistance. Based in part on a binding pocket analysis of both solved and unsolved protein structures by Sarkar and Brenk (2015), *PaPanK* was selected as a target for drug development [37]. A number of studies have found the *coaX* gene encoding *PaPanK* to be essential for *P. aeruginosa* [38][39][40]. Further, *coaX* from *P. aeruginosa* has been shown to rescue *coaBC* knockout *E. coli* [41]. Pantothenate kinase inhibitors have previously been developed against type I and type II isotypes, but such compounds have been shown to bind type III PanKs only weakly or not at all. In addition to the lack of amino acid identity, Guan et al. (2010) attribute this to the aforementioned closed substrate binding pockets of type III PanKs. This steric hindrance prevents known inhibitors of type I and type II PanKs, such as pantothenamides, from binding to the active site [42].

## 1.6 Fragment screening

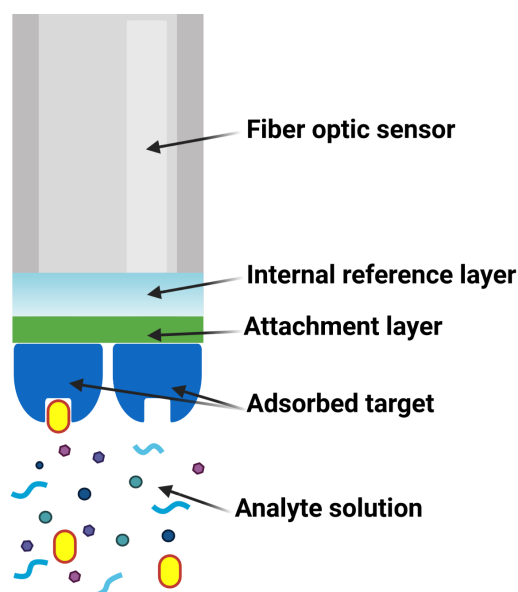
There exists no practical way to test all potential inhibitors in a laboratory setting given the existence of a vast number of compounds. Thus, diverse low molecular weight chemical libraries have been designed to probe binding pockets more efficiently than larger (~450 Da) drug-sized molecules. According to Edfeldt et al. (2011), testing a library of 1000 compounds under 200 Da can be as informative as a library consisting of 10 trillion drug-sized compounds [43]. The small size of these fragments reduces the chances of steric hindrance as well as excessive hydrophobic interactions from interfering with the screening process. While this means that smaller fragments can be used to effectively probe target sites and assess the abilities of different functional groups to interact with key residues, it also means that hits from a fragment screen tend to bind much more weakly than therapeutic drugs [44]. It is for this reason that

further refinement of initial fragment screening hits is often necessary to elicit stronger binding. Shuker et. al. (1996) demonstrated the additivity of changes in Gibb's free energy ( $\Delta G$ ) when two or more weakly-binding fragments are covalently linked, resulting in binding that is more energetically favorable and thus significantly stronger than either individual fragment alone. As long as the linked fragments can occupy adjacent spaces simultaneously, the small fragments can be combined into drug-sized molecules without the inefficient and costly process of screening drug-sized libraries [45]. Owing to the difficulty of linking fragments while retaining the binding modes of the original compounds and the difficulties of screening drug-sized compounds, it is increasingly common to gradually modify fragment screening hits through trial and error to discover drug-candidates (Figure 1.3) [46].



**Fig. 1.3:** Fragment-based drug design versus drug-sized, high throughput screening. (a) Large, drug-sized compounds (orange) are less efficient at probing the active site of the protein (blue) due to an increased likelihood of forming non-favorable interactions. Hit rates from screening such compounds are very low, but can have much higher affinities if the number of non-favorable interactions is limited. (b) Fragment-based drug design relies on screening of small compounds to identify scaffolds (1) which can be chemically altered or enlarged (2, 3) to achieve strong binding.

Identification of these weakly-binding compounds from a fragment library requires techniques which are both sensitive and, ideally, high-throughput. One such methodology is bio-layer interferometry (BLI). Using BLI, molecules of interest (including proteins) can be adsorbed onto sensor probes in various manners depending on the sensor surface coating before exposure to solvents containing molecules of interest (Figure 1.4). The fiber-optic sensors transmit white light through the pins down to the coated sensor tips, where it is reflected both at the interface of the fiber – bio-layer interface and the interface between the buffer and any bound molecules. The degree of interference from these two reflected beams of light depends on the added thickness of the bound molecules and can be used to detect the interactions between proteins and low molecular weight ligands. The signal contribution of non-specific binding can be mostly be accounted for by using a parallel set of reference sensors to which a quenching



**Fig. 1.4:** Bio-layer interferometry sensor visualization. The target of interest is adsorbed onto the attachment layer, which can consist of  $\text{Ni}^{2+}$ , streptavidin, or otherwise. The sensor is exposed to an analyte solution containing the compounds of interest, and light is passed through the fiber optic sensor. Binding to the adsorbed target changes the way light is reflected at the analyte solution - target interface, and results in a measurable change in interference with the light reflected by the internal reference layer. Figure created with Biorender.

compound is adsorbed. In the case of streptavidin-coated BLI sensors, this quencher can be either biotin or biocytin. In comparison to the more commonly used label-free technique of surface plasmon resonance (SPR), changes in the refractive index interfere less with BLI measurements [47]. Although BLI measurements with the Octet Red384 (ForteBio) are on the order of 10-fold less sensitive than with SPR instruments, it has been shown to be an effective fragment-screening tool for compounds  $>150$  Da. [47][48]. In this thesis, fragment screening experiments were conducted using the similar but smaller-scale Octet Red96 (ForteBio).

A virtual screen as conducted by Lee (2020) for potential inhibitors of *Pa*PanK identified primarily compounds with a carboxylate group [49]. In the simulated binding modes the carboxylate was seen to interact with Tyr92, Gly99, Arg102, and Thr180, similarly to the carboxylate of pantothenate (Figure 1.2). Therefore, these compounds plus an additional 305 compound library consisting solely of carboxylate-containing fragments were utilized in this work in the hopes of observing active-site specific binding.



## 1.7 X-ray crystallography

X-ray crystallography is an invaluable method for determining the 3D structures of proteins, in which crystallized proteins are subject to powerful X-ray beams in order to produce solvable diffraction patterns. This methodology requires that the molecule(s) of interest first be formed into orderly crystal structures with regularly arranged unit cells. The X-ray beams used are of similar wavelengths to short atomic bond lengths,  $\sim 1$  Å. The unit cell determines the scattering of the beams, while the layout of individual atoms within these unit cells determines the intensity of reflected beams [50]. Only diffracted beams which are in phase produce the constructive interference needed for detection. This only occurs when Bragg's Law (1.1), relating a whole number of wavelengths ( $n\lambda$ ), distance between diffraction points ( $d$ ), and beam angle relative to the lattice plane ( $\sin(\theta)$ ) is fulfilled [51].

$$n\lambda = 2d\sin(\theta) \tag{1.1}$$

In order to resolve these diffraction patterns into a 3D structure, the phase must also be determined as phase is a necessary component of the fourier synthesis used to map electron density [52]. As mentioned earlier, Se-Met *PaPanK* crystals have been previously diffracted both as an apo-enzyme (PDB 2F9T) and in complex with pantothenate (PDB 2F9W), with selenomethionine replacing all 6 methionine residues [35]. The presence of selenium in these residues produces anomalous, highly intense signals in the crystal diffraction patterns and is often utilized in solving the phase problem [53].

Dauter (1999) discusses the underlying principles of data collection from crystals when X-rays of a single wavelength are applied, as is typically performed for structural determination. Whether an atom reflects an X-ray beam depends on its position in an Ewald sphere of radius  $1/\lambda$  - only reciprocal lattice points along the edge of this sphere will satisfy Bragg's Law, resulting in a reflected beam which can be picked up by the detector and contribute to the overall diffraction pattern. At a single wavelength, the Ewald sphere remains constant and thus rotation of the crystal during exposure to the X-ray beam is necessary. Rotation of the crystal brings different reciprocal lattice points to the surface of the Ewald sphere, and with enough rotation enough

reflected beams can be measured to map the electron density of an entire protein. The amount of rotation necessary depends upon the crystal geometry, derived from the space group, as this determines which reciprocal lattice points will lay upon the edge of the Ewald sphere during beamline operation [54]. While crystals can be rotated in a helical manner or imaged serially from entirely separate crystals to reduce the impact of radiation damage, rotating about a single point is commonplace [55] [56].

The formation of protein crystals is itself a process of trial and error. The ultimate goal of such trials is to coax the protein of interest into super-saturated state that favors crystal nucleation, followed by growth [57]. The sitting-drop vapor diffusion method is widely used to produce protein crystals, where the protein of interest is added at a specific concentration to a drop of a given pH, precipitant, salt, and any other desired additives. The drop is sealed, sitting above a chamber containing the same precipitating buffer. By sealing this environment from outside air, vapor gradually diffuses and initially dehydrates the protein-precipitant drop, potentially leading to super-saturation and crystal formation before equilibrium is reached [58]. Each of the aforementioned drop contents affect the outcome, and the exact conditions required are unpredictable hence the typical need for many trials. If the protein is over-saturated, disordered aggregates or films can form which are not useful for structural determination. Undersaturation means that the solubilized state remains energetically favorable, and crystals can neither form nor grow. If conditions are energetically favorable for crystal growth, these crystals will continue to grow faster than they dissolve until reaching equilibrium, where crystal growth and dissolution rates are equal [59].

# Chapter 2

## Aims

The overarching goal of this work was to identify starting points in the drug discovery of new antibiotics which could be used to treat carbapenem-resistant *P. aeruginosa* infections by inhibiting *PaPanK*. BLI was used to screen a large number of compounds, both from fragment libraries and a structure-based virtual screen, in order to identify those which possessed the highest affinities for *PaPanK*. *PaPanK* needed to be biotinylated for use with this system, and expressed in large enough amounts for many assays to be performed. The development of methods to reliably obtain well-diffracting *PaPanK* crystals was also desired to support structure-based drug design efforts.



# Chapter 3

## Methods

### 3.1 Buffer preparation

The components of all non-purchased buffers were weighed out using a Sartorius Pro 32/34F Balance (Sartorius) for masses greater than 0.5 g, and an AB104-S (Mettler Toledo) scale for smaller amounts. Buffers were then pH adjusted using NaOH and HCl solutions using a magnetic stir bar and a 780 pH Meter (Metrohm). Buffers to be used on the ÄKTA Pure 25 (Cytiva) instrument were filtered using a vacuum apparatus with MF-Millipore™ Membrane Filters, 0.22 µm pore size (Merck).

### 3.2 SDS-PAGE

Visual evaluation of protein purity and size was performed via SDS-PAGE, typically using Any kD Mini-PROTEAN TGX Precast Protein Gels (Bio-Rad) and occasionally 4-20% Mini-PROTEIN TGX Precast Protein Gels (Bio-Rad) when Any kD gels were unavailable. A Precision Plus Protein™ Dual Xtra or Dual Color Standard (Bio-Rad) was used for size comparison in each gel. Samples were mixed with a 6X Laemmli loading dye [60] followed by heat-denaturation at 95 °C for 5 minutes prior to loading. A constant 100 V current was applied for 10 minutes, increasing to 150 V for an additional 20-30 minutes. Gels were stained per manufacturer instructions using either 25 mL InstantBlue Protein Stain (Expedeon) or PageBlue Protein Staining Solution

(Thermo Scientific). Imaging took place on a ChemiDoc XRS+ System (Bio-Rad) with exposure set to automatic and with no filters in place.

### 3.3 Plasmids utilized

All plasmids arrived pre-assembled and ready for transformation. 6HisTev*PaPanK* was contained in a pET-28a(+)-TEV (Genscript) vector as described in Lee (2020) [49]. 6HisAvi*PaPanK* was contained in a near-identical pET-28a(+)-TEV vector, with the addition of an AviTag [61] encoding amino acids GLNDIFEAQKIEWHE and a GS linker just prior to the *PaPanK* sequence. GST-BirA consisted of a glutathione-S-transferase affinity-tag linked to *E. coli* biotin ligase, contained in a pGEX-6P-1 vector donated by Petra Hänzelmann (University of Wuerzburg).

### 3.4 Transformation of *E. coli* with 6HisTEV*PaPanK* and 6HisAvi*PaPanK*

The early stages of production of 6HisTEV*PaPanK* and 6HisAvi*PaPanK* were identical except for the plasmid used. 1  $\mu\text{L}$  of 40 ng/ $\mu\text{L}$  of the appropriate plasmid was added to 100  $\mu\text{L}$  of C43(DE3) cells (Lusigen) on ice, and stirred with the pipette tip. Cells mixed with the plasmid were then incubated on ice for 30 minutes and subsequently heat shocked at 42 °C in a water bath for 45 seconds. Cells were then placed on ice for 2 minutes, and 950  $\mu\text{L}$  room temperature Recovery Medium for Expression (Sigma Aldrich) was added. Tubes of cells were placed on a Thermomixer Comfort (Eppendorf) shaker at 250 rpm for 1 hour at 37 °C. Agar plates were prepared using autoclaved LB agar (10 g tryptone, 5 g yeast extract, 5 g NaCl, 15 g agar per liter). LB-agar was microwaved until completely liquefied and allowed to cool to below 70 °C before the addition of 50  $\mu\text{g}/\text{mL}$  final kanamycin. Plates were poured roughly 1 centimeter deep and allowed to solidify, and 100  $\mu\text{L}$  of transformed cells were spread onto the plates in close proximity to an open flame. Plates were sealed with parafilm and left to incubate overnight at 37 °C.

### 3.5 Glycerol stock preparation

A single transformed bacterial colony was inoculated into 50 mL of autoclaved LB medium (10 g tryptone, 5g yeast extract, 5 g NaCl per liter) containing the plasmid-specific antibiotic by leaving a coated pipette tip in the LB medium. Inoculated LB medium was incubated at 37 °C overnight with 200 rpm shaking, and a 1:1 v/v mixture with autoclaved glycerol was prepared. 400  $\mu$ L aliquots were made before flash freezing with liquid nitrogen and storage at -80 °C.

### 3.6 Lysis and purification of recombinant *PaPanK* cell pellets

Following pelleting of recombinant *PaPanK* bacterial cultures, lysis of each pelleted 1 L of original culture was performed at 8 °C by the addition of 30 mL of lysis buffer (20 mM Tris HCl pH 7.9, 500 mM NaCl, 20mM Imidazole, 10% (v/v) glycerol, 1 tablet cOmplete, EDTA-Free Protease Inhibitor Cocktail Tablets (Roche) per 120 mL final solution volume). A magnetic stir bar was placed inside and mixing took place at 500 rpm for 25 minutes. Dnase I (Sigma Aldrich) was added after 25 minutes to a final concentration of 1 U/mL, and mixing continued for an additional 5 minutes. After mixing, resuspended pellets were transferred to 50 mL falcon tubes. The samples were then placed on ice and subjected to sonication at an amplitude of 20 W with 10 second intervals for 1 minute. After 1 minute of resting, samples were sonicated for an additional 1 minute using the same settings. Following sonication, samples were transferred to high-speed round bottom centrifuge tubes and spun down (15000 RCF, 15 minutes, 4 C). Supernatant was transferred into clean 50 mL tubes and then syringe-filtered using Whatman Filter Units 0.2  $\mu$ M (GE). Supernatant was then kept on ice before IMAC using an ÄKTA Pure 25 (Cytiva).

Binding buffer was prepared with 20 mM Tris HCl pH 7.9, 500 mM NaCl, 20mM Imidazole, 10% (v/v) glycerol, and an elution buffer was prepared with 20 mM Tris HCl pH 7.9, 500 mM NaCl, 500 mM Imidazole, 10% (v/v) glycerol. The ÄKTA Pure 25 flow rate was set to 5 mL/min and a HisTrap HP 5 mL (Cytiva) column was equilibrated with binding buffer. Filtered supernatant was then pumped onto the

column and washed with 5 CV binding buffer. A linear gradient elution transitioning from 100% binding buffer to 100% elution buffer was applied to the column, and eluant was collected in a Deep Well Megablock, 2.2 mL (Sarstedt) 96 well plate. UV absorption was measured by the built in UV meter, and protein peaks were identified through the UnicornEval (GE) software, using the appropriate extinction coefficient at 280 nm to determine yield and concentration. Protein fractions of interest were then combined and concentrated to a volume of 5 mL or less using Amicon Ultracel Filter Units (Millipore) with 10 kD size cutoffs, simultaneously buffer exchanging using gel filtration buffer (50mM Tris HCl pH 7.5, 150mM NaCl).

Protein was then subject to Size-Exclusion Chromatography (SEC) using a HiLoad 26/600 Superdex 75 pg (Cytiva) column equilibrated with gel filtration buffer. Concentrated protein samples were loaded onto a loop attached to the ÄKTA Pure 25 system and the column was washed with 1.5 CV of gel filtration buffer. Flow through was collected and the protein peaks were identified through the UnicornEval (GE) software. Samples were then further concentrated and if necessary, buffer exchanged using Amicon Ultracel Filter Units on a Centrifuge 5417R (Eppendorf) at 10,000 rpm until the desired volume was obtained. Concentration was then measured in duplicates or triplicates using the protein A280 function on a NanoDrop ND-1000 (Thermo Fischer Scientific).

Some 6HisTev*PaPanK* designated for use in crystallography experiments was subject to TEV protease cleavage and reverse His-Trap purification (Method 3.7). SEC was omitted in such cases.

### 3.7 6His-tag cleavage and reverse His-Trap purification

Following an initial His-trap purification of *PaPanK* containing a TEV site and a His-tag, samples intended for cleavage were buffer exchanged to cleavage buffer (25 mM Tris HCl pH 7.4, 150 mM NaCl). TEV protease was expressed and purified in-house and stored in 50 mM Tris-HCl pH 7.5, 1 mM EDTA, 5 M DTT. TEV protease was added at a 1 : 10 mass ratio to *PaPanK* in cleavage buffer and incubated overnight at 8 °C for 14-16 hours, rotating on an SB-3 Rotator (Stuart) set to 4 rpm. Samples were



then immediately subjected to reverse His-trap purification in order to separate any uncleaved protein and TEV protease from the desired cleaved protein.

A HisTrap HP 5 mL (Cytiva) column was equilibrated with binding buffer (20 mM Tris HCl pH 7.9, 500 mM NaCl, 20 mM imidazole, 10% (v/v) glycerol), and TEV protease treated samples were loaded using an ÄKTA Pure 25 (Cytiva). The column was then washed with 5 CV of binding buffer and the flow through was collected. Depending on the experiment, elution of any protein still bound to the column was performed isocratically with 100% elution buffer (20 mM Tris HCl pH 7.9, 500 mM NaCl, 500 mM imidazole, 10% (v/v) glycerol) or with a linear gradient transitioning from 100% binding buffer to 100% elution buffer over 5 CVs. Fractions which showed high UV absorbance were viewed using the UnicornEval (GE) software, and samples were assessed for cleavage and purity by running SDS-PAGE. Concentration was measured using a NanoDrop ND-1000, and 5% (v/v) glycerol was added prior to flash freezing with liquid nitrogen. Cleaved protein was then stored at -80 °C.

### 3.8 GST-BirA expression and purification

The GST-BirA needed for enzymatic biotinylation of *PaPanK* used in the course of this work resulted from the transformation of a pGEX-6P-1 vector provided by Petra Hänzelmann (University of Würzburg) into Rosetta(DE3)pLysS *E. coli* cells. Transformation, expression, and purification were performed as described by Fairhead and Howarth (2015) [61].

A modified version of this protocol was performed in this work but not utilized for enzymatic biotinylation of *PaPanK*, in which the same steps were followed until lysed cells were pelleted and supernatant was collected. Using an ÄKTA Pure 25 (Cytiva) system, this supernatant was loaded onto a GSTrap FF 5 mL column (Cytiva) equilibrated with PBS pH 7.3, 1mM DTT, 0.1% (v/v) Triton X-100. After equilibration, sample was loaded at a flow rate of 4mL/min. The column was washed with 5 CV of the same buffer at a 5 mL/min flow rate. Isocratic elution followed using 5 CV of 50 mM Tris HCl pH 8.0, 10mM reduced glutathione, 1mM DTT, at a flow rate of 5 mL/min and elution fractions were collected in a Deep Well Megablock, 2.2 mL (Sarstedt) 96 well plate.

### 3.9 Enzymatic biotinylation of 6HisAvi*PaPanK*

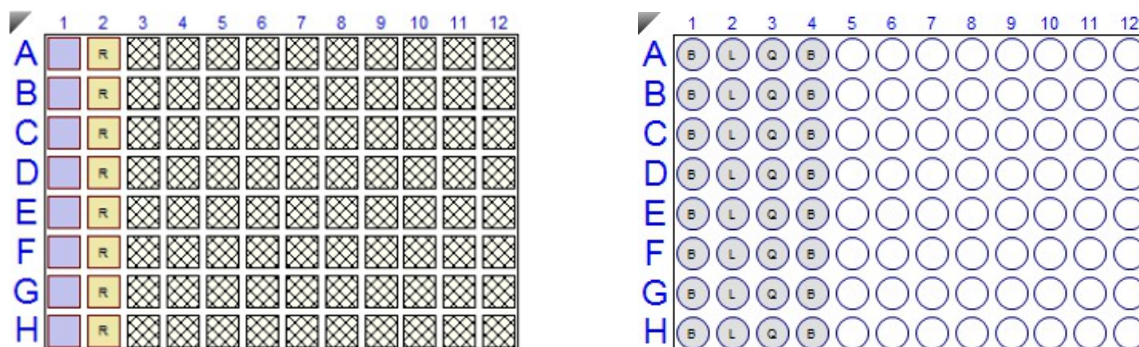
Site-specific enzymatic biotinylation was performed based upon the protocol by Fairhead and Howarth (2015) by incubating Avi-tagged *PaPanK* with *E. coli* biotin ligase (BirA). This BirA included a GST-tag in order to facilitate its removal after biotinylation.

952uL of 100uM 6HisAviPanK in gel filtration buffer (50mM Tris HCl pH 7.5, 150mM NaCl), 5uL 1M MgCl<sub>2</sub>, 50uL GST-BirA (1.45mg/mL stock), 3uL 50mM D-biotin (in DMSO), and 5uL 400mM ATP were added to a 1.5 mL Eppendorf tube. The mixture was incubated at 30 °C for one hour with 250 rpm shaking on a tabletop plate shaker. In order to remove GST-BirA, 0.1mL of glutathione 50% slurry was added to the mix and rotated for 1 hour at 80C. The resulting mixture was then centrifuged (4 °C, 2 minutes, 2k RCF). A Zeba Spin Desalting Column (Thermo Fisher) was washed twice with 2 mL of gel filtration buffer by centrifuging (4 °C, 2 minutes, 2k RCF). The desalting column was then transferred into a new eppendorf and the protein mix was added via pipette, avoiding the slurry mix. The protein was eluted by centrifuging (4 °C, 2 minutes, 2k RCF).

### 3.10 Bio-layer interferometry

Fragment screening of virtual screen hits and library fragments was performed using the Octet RED 96 Bio-Layer Interferometer (ForteBio). Octet Super-Streptavidin (SSA) Biosensors (ForteBio) were coupled with enzymatically biotinylated *PaPanK*, taking advantage of the especially strong affinity of streptavidin for biotin. In Black Sterile 96 Well F-Bottom Plates (Greiner), 200uL volumes were used per well to ensure that sensor tips were adequately immersed. SSA Biosensor racks were set up with two parallel columns of 8 sensors pre-soaked for 10 minutes in BLI running buffer (50mM Tris-HCl pH 7.5, 150 mM NaCl, 50 mM NH<sub>4</sub>Cl, 5% (v/v) DMSO). One set of sensors was designated for loading with biotinylated 6HisAvi*PaPanK* while a parallel reference set was designated for loading with biocytin (10 ug/mL). The instrument was set to 25oC.

During protein loading, sensors were serially exposed to a separate 96 well plate containing columns of BLI buffer, biotinylated *PaPanK*, and biocytin. First, sensors



(a) BLI sensor rack layout used in both protein loading and ligand analysis experiments. Two parallel columns of SSA biosensors are used, with those in light blue exposed to biotinylated *PaPanK* before quenching with biocytin. Yellow 'R' (reference) sensors skip the protein well and are only bound to biocytin.

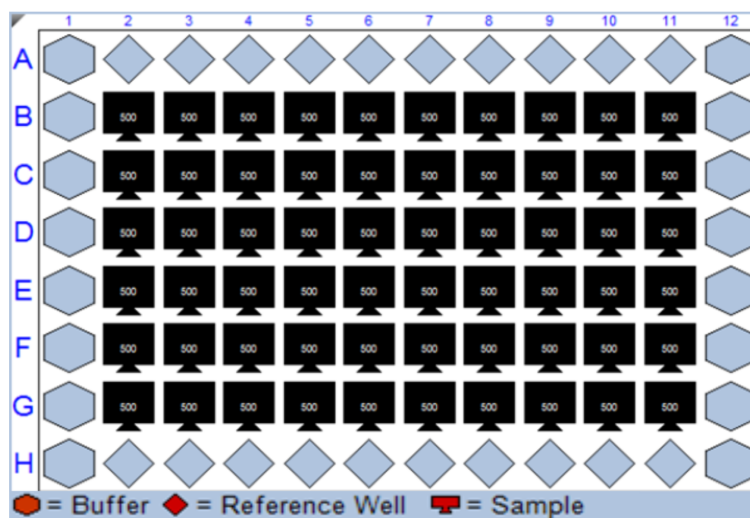
(b) BLI loading plate layout for preparing *PaPanK* sensors and biocytin reference sensors. Wells labelled 'B' contain BLI running buffer for equilibration, baseline, and wash steps. 'L' wells contain biotinylated 6HisAvi*PaPanK* which is not applied to the reference sensors. 'Q' contains the quencher, biocytin, applied to both the reference and protein-loaded sensors/

**Fig. 3.1:** Bio-layer interferometry sensor arrangement and well-layout for loading of sensors with the target protein and biocytin controls. The images were taken from the FortéBio Data Analysis 9.0 software.

were equilibrated for 300-1200 seconds with the BLI buffer, followed by a 60 second baseline step. Next, biotinylated *PaPanK* was bound to one column of sensors for 300 seconds. Next, sensors were transferred to biocytin for 60 seconds in order to quench any remaining unbound streptavidin molecules. A wash step proceeded for 60 seconds followed by a final 60 second baseline step. The reference sensor column was not exposed to biotinylated *PaPanK*.

### 3.10.1 Protein loading optimization

The determination of the concentration of biotinylated *PaPanK* necessary for eliciting a response  $> 5$  nm was performed by testing a range of concentrations from 12.5 to 250 ug/mL. Enzymatically biotinylated *PaPanK* (varying from 25 to 100 ug/mL) was bound to SSA biosensors in BLI running buffer (50mM Tris HCl pH 7.5, 150mM NaCl, 5% (v/v) DMSO, 50mM  $\text{NH}_4\text{Cl}$ ).



**Fig. 3.2:** Plate layout for fragment screening at a single concentration. When fragment screening at 2 concentrations with virtual screen compounds, alternating wells would be included at 2500  $\mu\text{M}$ . Rows A and H serve as reference-wells without added screening compounds to correct for sensor drift and background buffer signals. Columns 1 and 12 are also free of screening compounds and are used for the baseline and dissociation steps, respectively.

### 3.10.2 Compound binding

Fragment library compounds and virtual screen hits to be measured for  $K_D$  determination and *PaPanK* specificity arrived in 100% DMSO. For fragment screening of library compounds at a single concentration, a Mosquito HV was used to transfer 10  $\mu\text{L}$  of compounds into a 96 well plate. The same DMSO concentration was maintained in all wells. 190  $\mu\text{L}$  of DMSO-free BLI running buffer was added to all wells to which 10  $\mu\text{L}$  of DMSO-solubilized compounds were added. During fragment screening and hit validation of the carboxylic acid library and EU-OPENSOURCE (EU-OS) libraries, 0.01% (v/v) Triton X-100 was included in all wells. During hit validation, samples were pipetted manually and multiple 2-fold or 3-fold serial dilutions were performed for  $K_D$  determination. Water-solubilized samples such as pantothenate and ATP were first diluted from stocks into BLI running buffer before addition to 96 well plates, maintaining the 5% (v/v) final DMSO concentration.

#### BLI data processing

Data from the Octet RED 96 instrument were analyzed using the FortéBio Data Analysis 9.0 software. Compound identities, concentrations, and protein/biocytin

sensors and reference wells were defined. Double-referencing was selected to subtract the the biocytin and reference well responses from the protein responses. Baseline alignment was selected, Savitzky-Golay filtering (smoothing) was turned on, and the data were processed. For  $K_D$  determination, the built-in steady state model was applied using association step response values from 20-25 s or 50-55 s, depending on if the association step was 30 or 60 s.

For primary screening of the carboxylic acid and EU-OS libraries, biocytin and *PaPanK* sensor responses were only referenced against the buffer-only reference wells. The corrected biocytin sensor responses and *PaPanK* sensor responses were filtered within KNIME for each compound. A workflow in KNIME filtered these inputs by first excluding all samples with a biocytin response greater than 0.1 nm or less than -0.1 nm. The mean biocytin response  $\pm 1$  STD was calculated for the remaining compounds and any which exceeded the upper bound or were below the lower bound were excluded. Lastly, the mean *PaPanK* response - biocytin response + 1 STD was calculated and only compounds above this threshold were considered hits. These hits were subject to visual evaluation of the *PaPanK* response curves, with slow-binding compounds excluded.

### 3.11 *PaPanK* crystallization, ligand soaking, and structural determination

Crystallization of intact 6HisTev*PaPanK* following SEC purification was attempted by first concentrating the protein to 22 mg/mL in gel filtration buffer (50mM Tris HCl pH 7.5, 150mM NaCl) using Amicon 10k filter units. A Mosquito HV was used to prepare sitting drops on Triple Sitting Drop 96-well plates (TTP Labtech) with JCSG-plus, PACT Premier, Morpheus, and Ligand Friendly screens (Molecular Dimensions). 50  $\mu$ L of precipitant from each kit was manually pipetted into the reservoir wells. The Mosquito was used to create drops containing 100, 150, or 200 nL of reservoir-well precipitant and the amount of protein necessary to bring the final drop volume to 300 nL. One of each screen was prepared for storage at 20 °C and the plates were automatically imaged with a Rock Imager 182 (Formulatrix). Additional plates prepared with the

Morpheus, JCSG-plus, and Ligand Friendly screens were stored at 8 °C and manually examined using optical microscopy.

Aliquots of frozen TEV-cleaved *PaPanK* from a previous master's student were thawed and concentrated using 50 mM Tris-HCl pH 7.5 and optimized based on well E12 of the JCSG Plus screening plate (100 mM imidazole pH 8.0, 10% (w/v) PEG8000). Optimization of sitting drops based on a reproduction of this buffer was performed by varying pH, drop ratios, and input protein concentrations. In addition to varying well buffers and protein concentrations, 3 different drop ratios were used (100 nL protein:200 nL precipitant, 150 nL protein:150 nL precipitant, and 200 nL protein:100 nL precipitant) for a total of 288 conditions. The plate was stored and imaged at 20 °C. Further crystallization experiments utilized TEV-cleaved *PaPanK* expressed and purified during this thesis, with 100 mM imidazole pH 8.0, 8% (w/v) PEG8000, 150 nL protein:150 nL precipitant, and 17 mg/mL *PaPanK* in 50 mM Tris-HCl pH 7.5 stored at 20 °C.

### 3.11.1 Cryoprotectant selection

Protein crystals to be subjected to X-ray diffraction were cryoprotected prior to freezing in liquid nitrogen. Determination of cryoprotectants which did not dissolve or visually deform the crystals was performed using cryomixes (CM) from the CryoSol™ MD1-90 kit (Molecular Dimensions). These cryomixes contain several varying types and concentrations of cryoprotective compounds. Cryoprotective solutions of CM-1 through CM-9 were produced per manufacturer instructions using 50% (v/v) 2X crystallization precipitant, 40% (v/v) cryomix solution, and 10% (v/v) 10X crystallization buffer. Thereafter, 0.5 µL of the soaking solutions were added to various crystals and any visible changes to crystal structures were recorded after 21 hours. The predominantly used cryoprotectant was CM-3, which prior to dilution contained 25% (v/v) 2-methyl-2,4-pentanediol (MPD), 12.5% (v/v) 1,2-propanediol, 12.5% (v/v) DMSO, 12.5% (v/v) glycerol, 12.5% (v/v) ethylene glycol, and 25% (v/v) water.

### 3.11.2 Soaking of *PaPanK* crystals with ligands

Soaking of crystals in a cryoprotectant-ligand solution was performed after crystal formation in order to later diffract *PaPanK* crystals with ligands bound to the individual proteins making up the lattice. The chosen CM solutions from Method 3.11.1 were altered by making them with 32% (v/v) 3.125X crystallization precipitant and replacing the additional 18% volume with the ligand of interest. If the ligand was ATP or another nucleotide, up to 20 mM Mg<sup>2+</sup> and 80 mM K<sup>2+</sup> were included in the mixture. 0.5-0.8 μL of these soaking solutions were pipetted directly into crystal-containing drops and allowed to soak for up to 6 hours before fishing with loops and flash-freezing with liquid nitrogen.

### 3.11.3 Alternative soaking of *PaPanK* crystals with ligands

Alternatively, crystals were soaked using a cryomix-free solution which contained 32% (v/v) 3.125X crystallization precipitant, 10% (v/v) 10X crystallization buffer, and 36% (v/v) water, with the remaining 22% volume reserved for the ligand(s) of interest. Ligands dissolved in DMSO were limited to 8% (v/v) in the soaking solution. 0.8 μL of these soaking solutions were pipetted directly into the crystal-containing drops and allowed to soak for up to 6 hours. Thereafter, crystals were fished and briefly dipped into a drop of cryoprotective solution containing CM-3 (40% (v/v) CM-3, 32% (v/v) 3.125X crystallization precipitant, 10% (v/v) crystallization buffer, 18% (v/v) water) before flash-freezing with liquid nitrogen.

### 3.11.4 Structural determination and refinement of X-ray data

To produce molecular models from diffracted crystals of apo-enzyme *PaPanK* and *PaPanK*-Pantothenate, Se-Met *PaPanK* (PDB 2F9T) and Se-Met *PaPanK*-Pantothenate (PDB 2F9W) were used as reference models in the DIMPLE pipeline from CCP4i2. These models were subject to manual correction using Coot, while Refmac5 was used to improve agreement between the X-ray data and the molecular models. For modelling and refinement of the *PaPanK*-ATP and *PaPanK*-B1 diffraction data, the model produced by the *PaPanK* crystal diffracted during this work was used as the reference

model. Refinement of the *PaPanK*-ATP and *PaPanK*-B1 was not fully completed. The relevant data and statistics are shown in Appendix Table [A.1](#).

### **3.12 Work contributed by others**

The refinement of X-ray diffraction data was performed by Charis Georgiou. Vladyslav Yadrykhinsky produced the KNIME workflow utilized in fragment screening analysis. Khanh Kim Dao expressed and purified the TEV protease used for tag cleavage. Ludvik Olai Espeland expressed and purified the GST-BirA utilized during enzymatic biotinylation of 6HisAvi*PaPanK*.



# Chapter 4

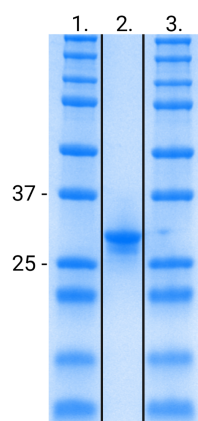
## Results

### 4.1 Expression and Purification of 6HisAvi*PaPanK*

One aim of this thesis is to identify small compounds with *PaPanK* specificity that could later be used to develop antibiotics. To screen large numbers of compounds, BLI was utilized. In order to obtain *PaPanK* which could be loaded onto the SSA sensors during BLI 6HisAvi*PaPanK* needed to be expressed, as it contains a 15 amino acid tag necessary for site-specific enzymatic biotinylation by *E. coli* Biotin Ligase (BirA). Following transformation into *E. coli* C43 (DE3) cells, a number of well-isolated colonies were observed growing on antibiotic-containing LB agar plates, which were successfully used for the preparation of glycerol stocks and IPTG-induced overexpression. After following the method described in “Lysis and purification” through to the His-trap column purification, a narrow peak was observed and collected in the gradient elution step at 40% elution buffer (Figure A.1), denoted as peak B. Based on the built-in UV meter and the  $A_{280}$  0.1% of 0.826, a total of 36.50 mg of 6HisAvi*PaPanK* was present in this peak.

The elution fractions from peak B from the His-Trap column were further processed as described in “Lysis and purification” and SEC was performed in order to isolate 6HisAvi*PaPanK* based on size (Figure A.2). Two small peaks followed immediately by a larger peak were observed, and this large peak C was suspected to contain the purified protein. Owing to the proximity of peaks A and B, the elution wells which

overlapped were not included when pooling the suspected protein fractions. According to the built-in UV meter and the 0.826  $A_{280}$  0.1%, this peak contained 15.01 mg of protein, equal to 7.50 mg of purified protein per liter of *E. coli* culture. The total protein yield according to Nanodrop was 14.93 mg, corresponding to a yield of 7.46 mg per liter of cultured *E. coli*. SDS-PAGE was performed in order to assess the size and purity of the protein (Figure 4.1). One prominent band appeared in Lane 2



**Fig. 4.1:** SDS-PAGE after staining with PageBlue. Lanes 1 and 3 contain Precision Plus Dual Color Protein Standards, with the labelled bands denoting size in kDa. Lane 2 contains the pooled contents from SEC elution suspected to contain 6HisAviPaPanK

between the 25 and 37 kDa ladder markers, consistent with the expected 31.4 kDa size of 6HisAviPaPanK. With the size confirmed, the protein was deemed pure enough to attempt enzymatic biotinylation.

In order to uniformly biotinylate the Avidin-tagged PaPanK for use with BLI, enzymatic biotinylation was performed using *E. coli* Biotin Ligase (BirA) as described in Method 3.9. The success or failure of biotinylation was later assessed by loading to the SSA sensors during BLI.

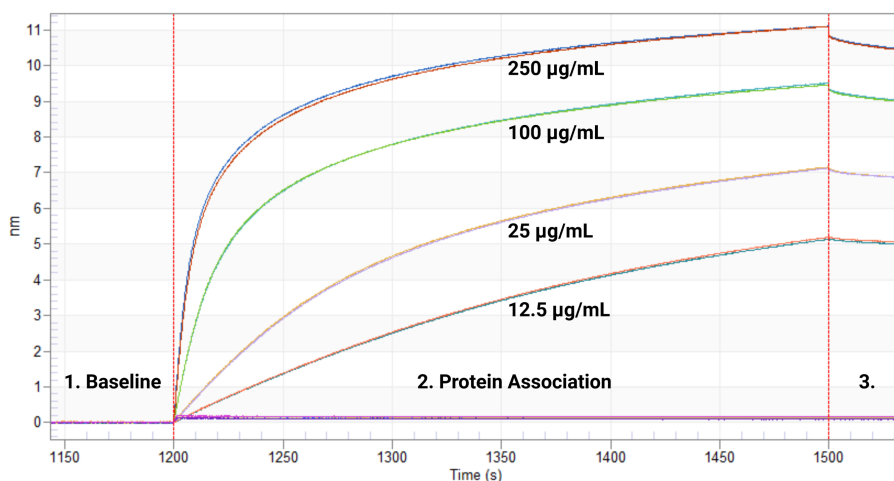
## 4.2 BLI with PaPanK

### 4.2.1 Protein loading optimization

Assessment of the binding of small molecules to the biotinylated 6HisAviPaPanK first requires coating of the BLI sensors with the biotinylated protein of interest. SSA sensors

were used due to their higher binding site density compared to the manufacturer's traditional streptavidin sensors, thus allowing for more biotinylated protein to be bound. A balance between high sensor binding and minimal protein usage was desired, and a range of protein concentrations were tested as described in Method 3.10.1 At each concentration, protein was loaded onto SSA sensors as duplicates except for 100  $\mu\text{g}/\text{mL}$ , which was loaded onto only one sensor to accommodate a control sensor exposed only to buffer.

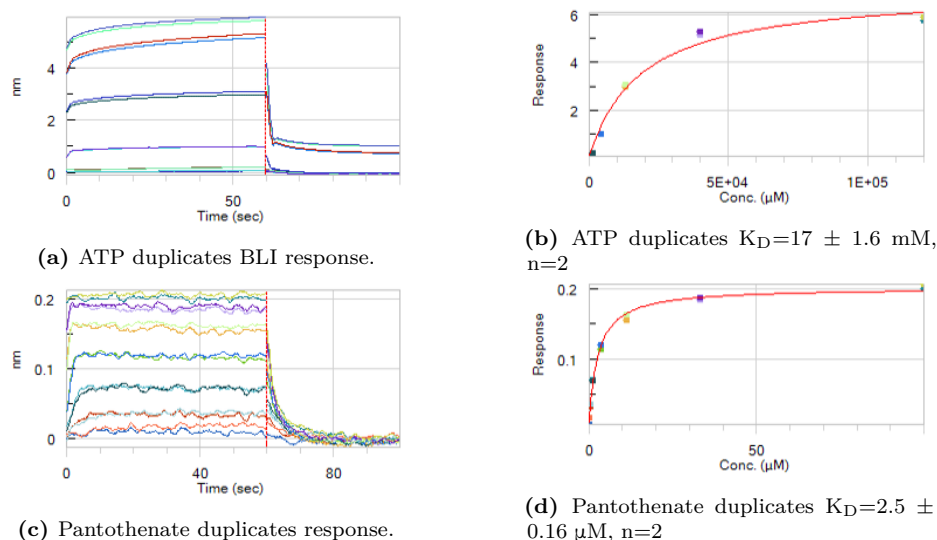
After 300 seconds of exposure to various concentrations of biotinylated 6HisAvi*PaPanK*, response curves were produced (Figure 4.2). Responses increased with increased protein concentration and no-protein controls elicited minimal responses. Protein concentrations  $\geq 25$   $\mu\text{g}/\text{mL}$  were deemed suitable for use in future binding analysis experiments, and 25 or 100  $\mu\text{g}/\text{mL}$  biotinylated 6HisAvi*PaPanK* was used for loading in all subsequent BLI experiments.



**Fig. 4.2:** Protein loading confirmation of biotinylated 6HisAvi*PaPanK* onto streptavidin sensors for BLI experiments. 1) Baseline step as a point of reference during analysis 2) Protein association step with variable *PaPanK* concentrations. 3) Dissociation step where sensors were transferred to protein-free buffer wells. The flat, nearly 0 nm response line comes from a sensor not exposed to biotinylated 6HisAvi*PaPanK*.

### 4.2.2 Establishment of ATP and pantothenate binding

In order to determine the suitability of BLI for identifying the binding of ligands to *PaPanK*,  $K_D$  determination of ATP and pantothenate was performed. As described in Introduction 1.4, ATP and pantothenate have been previously shown to bind to



**Fig. 4.3:** Combined duplicates of ATP and pantothenate BLI responses and calculated  $K_D$  values showing saturation with increased ligand concentrations. Figures (a) and (c) show double referenced responses for ATP and pantothenate through association steps (0-60 s) followed by dissociation (denoted by vertical red line). Figures (b) and (d) show the fitted lines of duplicate ligand concentrations versus response values from steady state analysis. Error was calculated from the individual steady state curves.

*PaPanK* with respective  $K_M$  values of 7.3 mM and 5  $\mu$ M. Protein was loaded onto the sensors using 100  $\mu$ g/mL and reached consistent responses of 10 nm. 5% (v/v) DMSO was included in all baseline, association, and dissociation wells since this would be present during the subsequent screening of the DMSO-solubilized fragments.  $K_D$  determination was performed according to Method 3.10.2 with ATP and pantothenate tested in duplicates, with each replicate row being exposed to different protein-loaded sensors and biocytin-loaded controls. 7 2-fold dilutions were performed, beginning from 100  $\mu$ M pantothenate and 120 mM ATP.

In each duplicate set, the sensor response to each ligand increased as concentration increased before flattening out as the maximum concentrations were approached, and response values and fit curves were generated for ATP and pantothenate (Figure 4.3). Duplicate response values for each ligand were highly consistent and produced well-fitting steady state curves, with R-squared values of 0.994 for pantothenate (Figure 4.3 d) and 0.987 for ATP (Figure 4.3 b). The experimental  $K_D$  values when combining duplicates of pantothenate and ATP were 2.5  $\mu$ M and 17 mM respectively, versus the literature  $K_M$  values of 5  $\mu$ M and 7.3 mM. While the  $K_D$  value of pantothenate indicated much higher affinity for the biotinylated *PaPanK* sensors than ATP, the ATP

responses were notably much larger. ATP also did not appear to dissociate completely at concentrations over 40 mM (Figure 4.3 a). Based on the behavior of pantothenate, it was selected for continued use as a positive control during forthcoming fragment screening and hit validation experiments.

### 4.2.3 Testing of virtual screening hits

A virtual screen conducted by Lee (2020) identified a number of compounds which were predicted to bind to the active site of *PaPanK*, as described in 1.6. 17 of these compounds were purchased and subsequently solubilized in DMSO for use in BLI. The decision was made to conduct fragment screening of these compounds twice at different pH values (pH 6.0 and pH 7.5) in case pH 7.5 did not support the protonation states necessary for certain compounds to bind. The pH 6.0 buffer was identical except for the use of 50 mM Bis-Tris Methane pH 6.0 in lieu of the Tris-HCl used in previous BLI experiments (50mM Tris HCl pH 7.5, 150mM NaCl, 5% DMSO, 50mM NH<sub>4</sub>Cl). Sample sensors were loaded with 25  $\mu$ g/mL *PaPanK* and reached response levels of 7 nm at pH 7.5 and 10 nm at pH 6.0. The 17 virtual screen compounds were prepared at concentrations of 2500  $\mu$ M and 500  $\mu$ M. 60 s baseline, 60 s association, and 60 s dissociation steps were used and the screening was conducted at the two pH values.

Sample ID	500 $\mu$ M Response (nm) pH 7.5	500 $\mu$ M Response (nm) pH 6.0	2500 $\mu$ M Response (nm) pH 7.5	2500 $\mu$ M Response (nm) pH 6.0
a1	0.113	0.0087	0.1675	0.0277
a2	-0.0373	-0.0225	-0.0729	-0.0593
a3	0.0031	0.0916	0.0811	0.2599
a4	-0.0078	0.0078	-0.0272	0.0075
b1	0.1541	0.0272	0.2732	0.1491
b2	-0.0323	-0.0216	-0.0204	-0.0754
b3	0.0069	0.0006	0.0943	0.0126
b4	0.0234	-0.0026	0.0365	-0.0044
c1	0.0272	0.0042	0.0082	0.0181
c2	0.0213	-0.0056	0.1111	-0.0162
c3	-0.036	0.0193	-0.0523	0.0976
d1	0.0781	0.0162	0.1322	0.0293
d2	-0.0149	0.0287	-0.0353	0.1608
d3	0.0122	0.0328	0.0522	0.1409
e1	-0.0341	0.0246	-0.0231	0.1383
e2	0.025	-0.0361	0.1342	-0.0829
e3	0.0197	-0.0063	0.0465	-0.0467

**Table 4.1:** Double-referenced BLI responses following 2-point screening of 17 virtual screen compounds at pH 6.0 and pH 7.5. Compounds with highlighted response values were designated for further hit validation based on their response values.

Responses from the *Pa*PanK sensors and the biocytin sensors were recorded and double referencing was applied using the FortéBio Data Analysis 9.0 software according to Method 3.10.2. These corrected response values were reviewed and compounds were selected for further hit validation based on response intensity (Table 4.1). Notably, response values varied considerably depending on the pH, with only two compounds showing responses  $> 0.05$  nm at both pH 6.0 and pH 7.5 (Table 4.1 b1, d3). Several compounds showed negative responses at one or more pH, with a few becoming significantly more negative with increased concentration (Table 4.1, a2, a4, b2, c2, c3, d2, e2, e3). 11 compounds showed response values greater than the chosen cutoff of 0.05 nm at 2500  $\mu$ M in at least one pH value. One of these compounds (Table 4.1, a3) was subsequently excluded based on a linearly increasing response during association and near-complete lack of dissociation from the sensor. The remaining compounds (Table 4.1, a1, b1, b3, c2, c3, d1, d2, d3, e1, e2) were considered candidates for further hit validation.

### Virtual screen hit validation

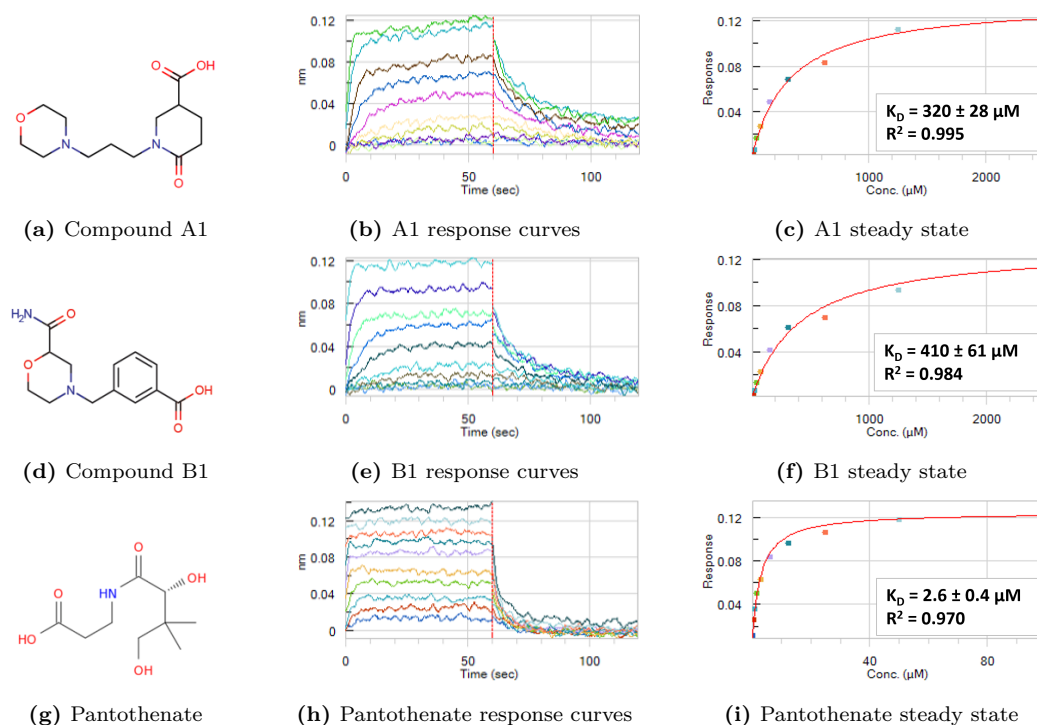
In order to determine the  $K_D$  of the 10 virtual screen compounds narrowed down in Section 4.2.3, 10 2-fold serial dilutions from 2500  $\mu\text{M}$  to 4.88  $\mu\text{M}$  were prepared prior to BLI. While there was little overlap in responses at different pH values from the virtual screening compounds, all 10 compounds were hit-validated at both pH 6.0 and pH 7.5 owing to the small amount of scale-up required. *PaPanK* was loaded at 25  $\mu\text{g}/\text{mL}$  and the *PaPanK* exposed sensors reached 11 nm response at pH 6.0 and 7 nm response at pH 7.5. This pH or buffer-dependent variation in protein loading response was also observed in Section 4.2.3. The serially diluted compounds were then screened with 60 s baseline, 60 s association, and 60 s dissociation steps.

Sample ID	pH 6.0 2500uM Response (nm)	pH 6.0 $K_D$ (M)	pH 6.0 $R^2$	pH 7.5 2500uM Response (nm)	pH 7.5 $K_D$ (M)	pH 7.5 $R^2$
a1	0.0292	5.70E-03	0.9875	0.1228	3.20E-04	0.995
b1	0.0872	1.50E-02	0.9961	0.1206	4.10E-04	0.9843
b3	-0.0011	3.50E+11	0.649	0.0098	4.20E-04	0.6755
c2	-0.013	<1.0E-12	0	-0.0362	<1.0E-12	0
c3	0.0432	<1.0E-12	0	0.0143	1.60E-03	0.7056
d1	0.0216	1.40E-02	0.9921	0.0896	9.40E-04	0.9828
d2	0.0551	1.20E+12	0.9406	0.079	8.50E-03	0.9915
d3	0.0599	<1.0E-12	0	0.1383	1.80E-02	0.9885
e1	0.0636	5.60E-02	0.967	-0.011	-	-
e2	-0.0443	2.00E-05	1	0.093	9.90E-04	0.9625

**Table 4.2:** Doubled-referenced BLI responses and calculated  $K_D$  values with steady state graph curve fit  $R^2$  following 10-point dilution of 10 virtual screen compounds. Both pH 6.0 and pH 7.5 results are shown, and the 2 hits considered most promising are highlighted in blue.

Of these 10 compounds screen at 2 varying pH values, the majority showed either low response values ( $< 0.05$  nm), millimolar or higher  $K_D$  values, or poor steady state curve fits. Based on this, no compounds from pH 6.0 were considered for further analysis. Compounds D3 and E2 at pH 7.5 showed acceptable response at 2500  $\mu\text{M}$  and showed high micromolar  $K_D$  values, but were excluded from further evaluation based on their association behaviors. The 2 compounds considered to have good binding behavior at pH 7.5 were A1 ( $K_D=320 \pm 26$   $\mu\text{M}$ ) and B1 ( $K_D=410 \pm 61$   $\mu\text{M}$ ), the structures and binding behaviors of which are shown in Figure 4.4.

Whether or not these compounds competed with the binding of pantothenate was further investigated. In a separate BLI test, compound A1 was maintained at a constant



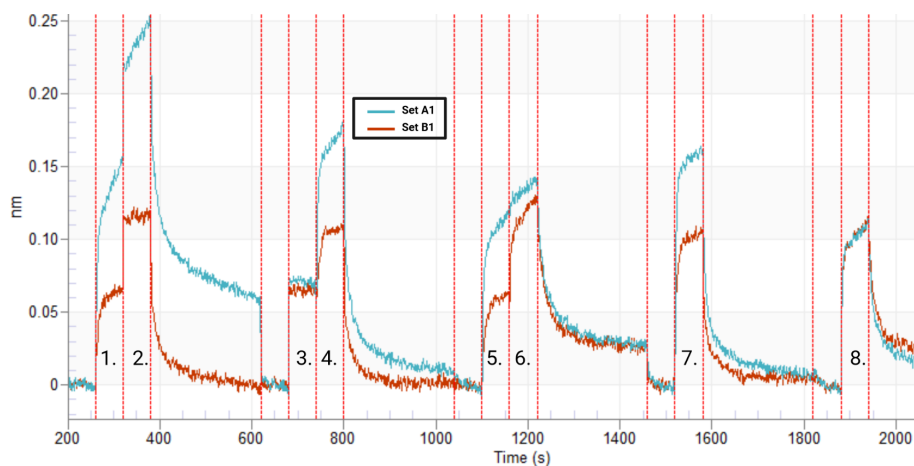
**Fig. 4.4:** Structures and hit validation of the most promising compounds tested from the virtual screen (a-f) and the pantothenate control included in the same test (g-i).  $K_D$  and  $R^2$  of these compounds are shown in the steady state graphs (c, f, i)

concentration of  $1250 \mu\text{M}$  as pantothenate was serially diluted 2-fold from  $50 \mu\text{M}$  to  $0.19 \mu\text{M}$ . Triplicate measurements of the  $K_D$  value of pantothenate in the presence of constant A1 were performed, including duplicate measurements in the absence of A1. Biotinylated 6HisAviPaPanK was loaded to a response of  $7 \text{ nm}$  at pH 7, and the compound responses were measured with 60 s baseline, 60 s association, and 60 s dissociation steps. The mean  $K_D$  value of pantothenate in the absence of A1 ( $3.1 \pm 0.18 \mu\text{M}$ ,  $n=3$ ) was not increased in the presence of A1, but found to decrease ( $1.6 \pm 0.29 \mu\text{M}$ ,  $n=2$ ). The test was not repeated on compound B1 due to lack of available compound.

An alternate test dependent only on response values was performed with exposures of sensors with biotinylated 6HisAviPaPanK and biocytin sensors to combinations of A1, B1, and pantothenate. The goal was to determine whether the binding of each compound was additive, partially additive, or inhibitory. Without dissociation in between certain association steps, the sensors showed the differing behaviors of A1 and B1 (Figure 4.5). The main takeaway from this graph is that after binding of a single virtual screen hit (1.), the following association step (2.) with pantothenate added varying degrees of pantothenate response (3.) to the hit response (1.). in the case of



compound A1 the responses appeared completely additive, with the full responses of both pantothenate alone and A1 alone present when they were combined. In the case of compound B1 the responses were only partially additive, and thus B1 was tentatively considered to have at least some specificity for the pantothenate binding site. A1 was removed from further consideration, while structural determination of the binding mode via X-ray crystallography was later attempted for compound B1.



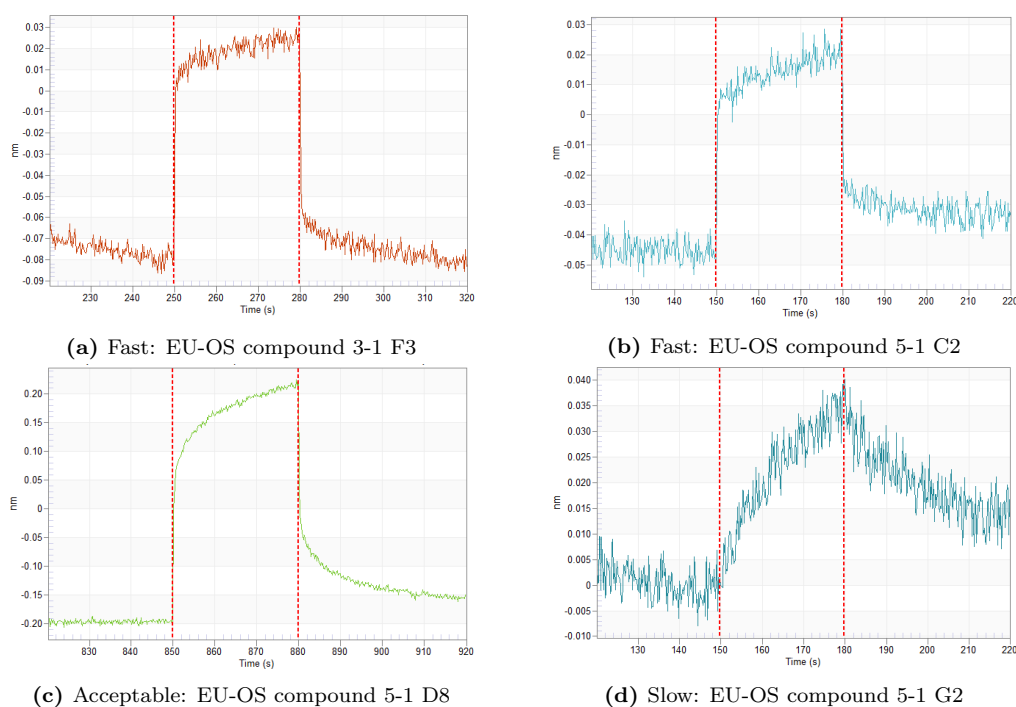
**Fig. 4.5:** BLI pantothenate inhibition testing of compounds A1 and B1, with association steps labelled. 1) 312.5  $\mu\text{M}$  of a single hit compound. 2) 312.5  $\mu\text{M}$  hit compound + 25  $\mu\text{M}$  pantothenate. 3)  $\mu\text{M}$  pantothenate. 4) 25  $\mu\text{M}$  pantothenate + 312.5  $\mu\text{M}$  of a single hit compound. 5) 312.5  $\mu\text{M}$  of a single hit compound. 6) 312.5  $\mu\text{M}$  of both A1 and B1. 7) 25  $\mu\text{M}$  pantothenate + 312.5  $\mu\text{M}$  of a single hit compound. 8) 312.5  $\mu\text{M}$  of both A1 and B1.

#### 4.2.4 Fragment screen of EU-OS library

In order to increase the chances of finding potential *PaPanK* inhibitors beyond reliance on virtual screening compounds, two separate fragment screening libraries were obtained. Each compound would be assessed by BLI to determine which compounds appeared to bind to *PaPanK*, while later refinement based on binding responses of the protein and biocytin reference sensors would be used to narrow the results prior to a secondary hit-validation experiment. The first library utilized was a 1047 sample subset of compounds curated by EU-OPENSREEN and designed for efficient fragment screening. Prior to screening of these compounds biotinylated 6HisAvi*PaPanK* (25  $\mu\text{g}/\text{mL}$ ) was bound to SSA sensors alongside biocytin (10  $\mu\text{g}/\text{mL}$ ) controls for detection of nonspecific binding as described in Method 3.10. Biotinylated 6HisAvi*PaPanK* response upon loading to sensors was consistent across all sensors to a level of 6 nm and the fragment

screening proceeded. Each of the EU-OS compounds was screened at single 500  $\mu\text{M}$  concentrations for the detection of binding responses, with 30 s baseline, 30 s association, and 40 s dissociation steps.

The biocytin sensor responses and *PaPanK* sensor responses were filtered to identify hits using a workflow (Method 3.10.2). Initial filtering by excluding compounds with responses  $> 0.1$  nm or  $< -0.1$  nm for biocytin sensors excluded 162 compounds. Next, the mean biocytin response  $\pm 1\text{STD}$  was calculated for the remaining compounds ( $0.0158 \pm 0.0308$ ) and a further 241 compounds with responses above this upper limit and below this lower limit were excluded. In the final step the mean *PaPanK* response minus biocytin response,  $+ 1$  STD was calculated ( $0.0096 + 0.0256$ ) and filtered out all compounds with responses  $< 0.027$  nm. This step filtered out 565 compounds.



**Fig. 4.6:** Raw BLI data for *PaPanK* sensors from the 500  $\mu\text{M}$  concentration screening of the EU-OS library, with examples of association/dissociation curves visually evaluated after hit processing. Near-instant binding to the *PaPanK* sensors was desired. The vertical red dotted lines separate baseline, association, and dissociation steps. (a) and (b) show compounds which were visually determined to have fast association and dissociation rates. The binding of (c) and 19 other compounds with similar behaviors were deemed acceptably fast. The binding of (d) and 41 other compounds with similar behaviors were deemed slow and excluded.

After all hit processing, 79 compounds (7.55% of evaluated compounds) were identified as candidates for secondary screening. The binding profiles of these 79 compounds

were evaluated visually in order to exclude compounds which associated or dissociated in slow, roughly linear manners (Figure 4.6 d). 42 compounds were excluded based on this criteria. Of the remaining compounds, 17 appeared to have unambiguously fast on/off binding curves (Figure 4.6 a, b), while a further 20 appeared to bind more slowly but not poorly enough to be excluded entirely (Figure 4.6 c). In total, 37 compounds were designated for further hit validation.

### EU-OS hit validation

The 37 EU-OS compounds filtered both by KNIME and by visual evaluation were subject to further testing and  $K_D$  determination using BLI and 5 serial 2-fold dilutions from 500  $\mu\text{M}$  to 31.3  $\mu\text{M}$ . Pantothenate positive controls were serially diluted in the same manner but from 25  $\mu\text{M}$  to 1.56  $\mu\text{M}$  and were present during the hit validation of each compound. DMSO from the hit-picking plate from which compounds were manually transferred was used to prepare the buffer-only reference wells and pantothenate positive controls. 30 s baseline, 30 s association, and 40 s dissociation steps were used for all compounds except for EUOS-hits 33-37, where the dissociation step was accidentally omitted. EUOS-hit 7 was not tested due to accidental omission.

The resulting response values and calculated fitted curves did not reveal any clear hits (Table 4.3). The pantothenate positive controls used on each plate appeared similar to those shown previously (Figure 4.3c). 27/36 tested EU-OS compounds showed response values  $< 0.027$  nm, the KNIME cutoff established in Section 4.2.4.  $K_D$  values for compounds above this response threshold tended to be in the high millimolar or molar range, or fit only moderately, weakly, or not at all. EUOS-hit 2 appeared acceptable, but given its low maximum response it was not selected for further evaluation.

EUOS Hit ID	Max concentration ( $\mu\text{M}$ )	Max Response (nm)	$K_D$	$R^2$
1	500	-0.0128	<1.0E-12	0
2	500	0.0289	7.20E-04	0.9778
3	500	0.0192	1.50E-04	0.7394
4	500	0.0287	1.80E-04	0.9239
5	500	0.0009	6.20E-06	0.004
6	500	0.0097	7.90E+10	0.9952
8	500	0.014	2.10E-03	0.9586
9	500	0.0225	<1.0E-12	0
10	500	0.0152	1.50E-02	0.96
11	500	0.0236	2.70E-04	0.9483
12	500	0.0003	<1.0E-12	0
13	500	0.0119	8.70E-05	0.7975
14	500	-0.0042	<1.0E-12	0
15	500	0.0151	7.50E-05	0.8867
16	500	0.034	<1.0E-12	0
17	500	0.0244	8.90E-05	0.9121
18	500	-0.0023	<1.0E-12	0
19	500	0.0091	7.70E+10	0.908
20	500	0.0017	<1.0E-12	0
21	500	0.0541	4.40E+11	0.8995
22	500	0.1368	<1.0E-12	0
23	500	0.0148	<1.0E-12	0
24	500	0.0601	9.30E+09	0.9941
25	500	0.0082	<1.0E-12	0
26	500	0.0518	1.30E-04	0.4456
27	500	0.2292	2.60E+11	0.9328
28	500	0.0275	2.70E-04	0.7549
29	500	-0.0112	<1.0E-12	0
30	500	0.0262	5.30E-04	0.9441
31	500	0.0024	<1.0E-12	0
32	500	0.0196	3.30E+11	0.8785
33	500	0.0034	4.30E-05	0.6859
34	500	0.0342	1.60E-03	0.9399
35	500	-0.3363	-	-
36	500	0.0245	4.50E-04	0.9864
37	500	0.014	5.60E-04	0.8804

**Table 4.3:** Results from hit validation of 36 EU-OS compounds.  $K_D$  values and  $R^2$  for BLI response fits to the steady state graph used for  $K_D$  determination are shown.

### 4.2.5 Fragment screen of carboxylic acid library

A diverse library of 305 carboxylic acids was ordered given the importance of this functional group in the interaction of pantothenate with several *PaPanK* active site residues as described in Introduction 1.6. Three of the 305 compounds posed solubility issues and thus the library which arrived contained 302 DMSO-solubilized compounds for testing. Prior to screening of these compounds biotinylated 6HisAvi*PaPanK* (25  $\mu\text{g}/\text{mL}$ ) was bound to SSA sensors alongside biocytin (10  $\mu\text{g}/\text{mL}$ ) controls for detection of nonspecific binding described in Method 3.10. Biotinylated 6HisAvi*PaPanK* response upon loading to sensors was consistent across all sensors to a level of 6 nm and the fragment screening proceeded. The 302 carboxylic acids were screened at single 500  $\mu\text{M}$  concentrations for the detection of binding responses, with 30 s baseline, 30 s association, and 40 s dissociation steps.

The responses for the sensors were processed through KNIME in the same way as described for the EU-OS library. The mean biocytin response  $\pm 1$  STD was calculated ( $-0.004 \pm 0.018$  nm) and compounds above this upper limit and below this lower limit were excluded. Calculation of the mean *PaPanK* response - biocytin response, + 1STD ( $0.012 + 0.039$ ) filtered out all compounds with responses  $< 0.051$  nm.

After all hit processing, 9 compounds (2.98% of evaluated compounds) were identified as candidates for secondary screening. The binding profiles of these 9 compounds were evaluated visually in order to exclude compounds which associated or dissociated in slow, linear rather than logarithmic manners or showed otherwise irregular behavior. 3 of the 9 compounds were excluded based on this criteria. Likewise, compounds which were narrowly filtered out by the KNIME workflow were visually evaluated and 4 such compounds were manually included. In total, 10 compounds were designated for further hit validation.

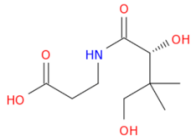
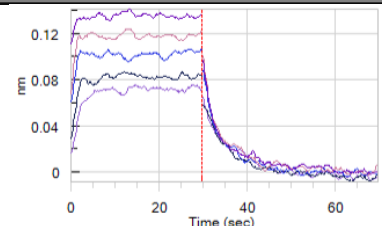
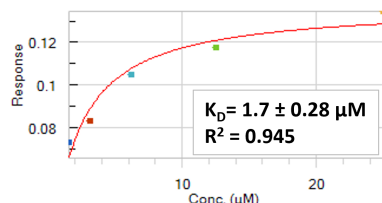
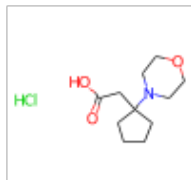
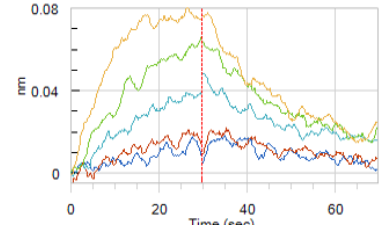
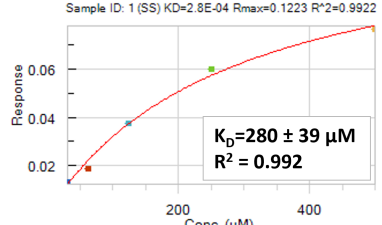
#### Hit validation of carboxylic acids

As was performed for the EU-OS hits, the  $K_D$  was calculated for the carboxylic acid hits in order to identify compounds with the highest *PaPanK* affinity. DMSO from the hit-picking plate from which compounds were manually transferred was used to prepare the buffer-only reference wells. A hit validation experiment using 5 2-fold

serial dilutions of the compounds selected from Section 4.2.5 was performed, with concentrations ranging from 31.3  $\mu\text{M}$  up to 500  $\mu\text{M}$ .

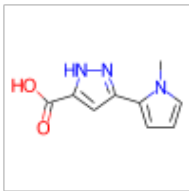
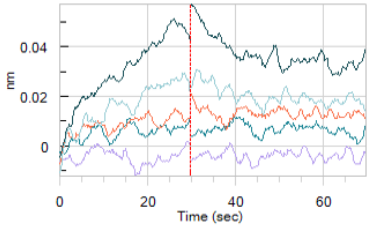
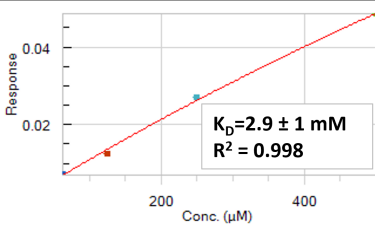
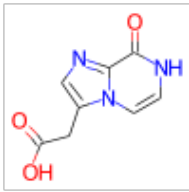
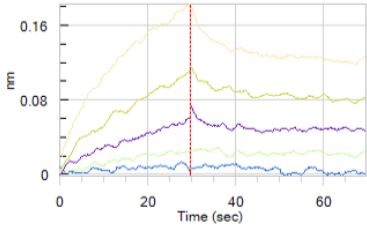
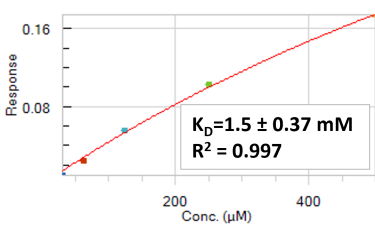
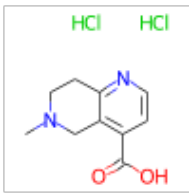
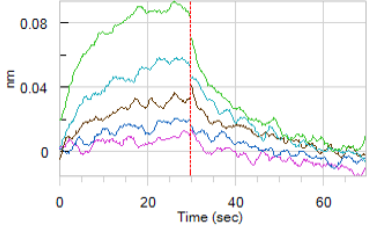
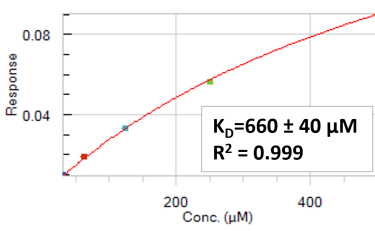
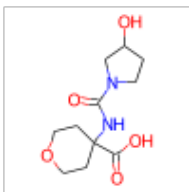
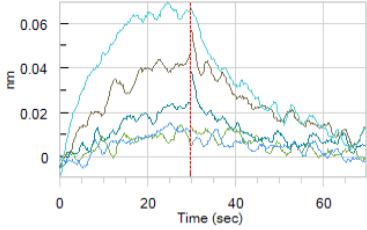
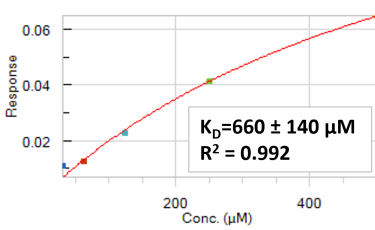
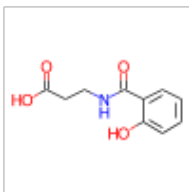
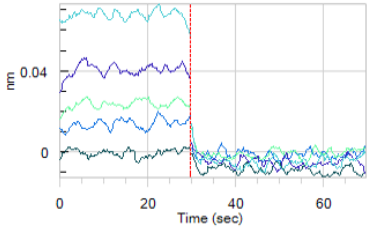
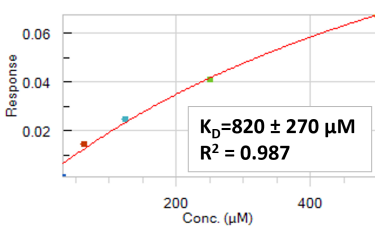
Response was measured, steady state analysis was carried out, and structures of the acid compounds were visualized (Table 4.4). The pantothenate control bound well to the sensor, with a calculated  $K_D$  of 1.7  $\mu\text{M}$ , similar to the previously determined value (Figure 4.3c). Hits 9 and 10 of the carboxylic acid library (Table 4.4) showed responses lower than the 0.051 nm threshold utilized previously (Section 4.2.5) and individual responses did not fit well to the steady state curve used to calculate  $K_D$ . Acid hits 2 and 3 (Table 4.4) were excluded as they did not appear to bind quickly to or dissociate well from the sensors and provided millimolar  $K_D$  values. Carboxylic acid hits 1, 4, 5, 6, 7, and 8 all produced sub-millimolar dissociation constants, with hit 7 (Table 4.4) showing the greatest affinity with a  $K_D$  value of 110  $\mu\text{M}$ . These binding responses and association/dissociation behaviors were also deemed acceptable.

**Table 4.4:** Resulting binding curves and steady state responses used for  $K_D$  determination of carboxylic acid hit validation compounds. The RDkit rendered structures are included. 'Hit' refers to the carboxylic acid compounds selected for hit-validation following Results 4.2.5. Column 'Binding' shows the double-referenced responses for each serially diluted compound. Column 'Steady State' shows the fitted curves used to calculate  $K_D$ . 'Pant' under the 'Hit' column refers to pantothenate, used as a positive control.

Hit	Structure	Binding	Steady State
Pant			
1			

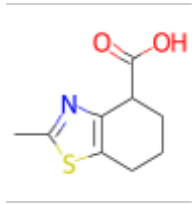
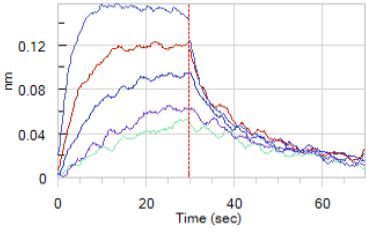
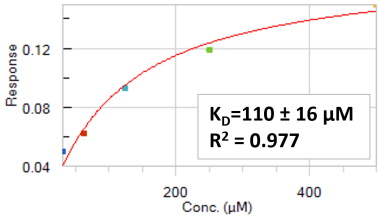
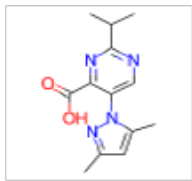
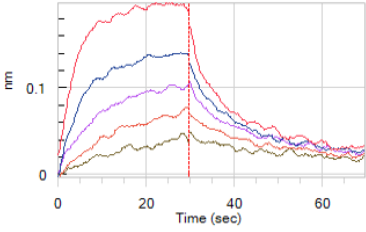
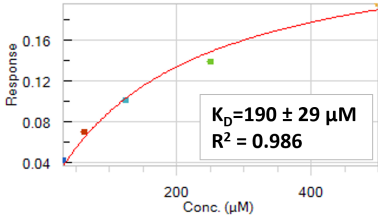
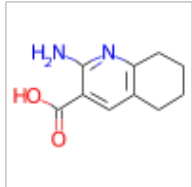
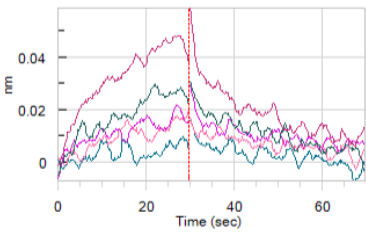
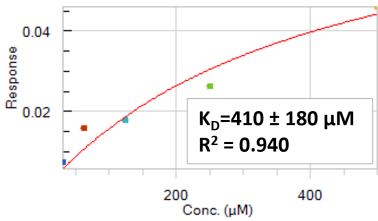
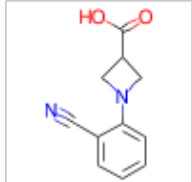
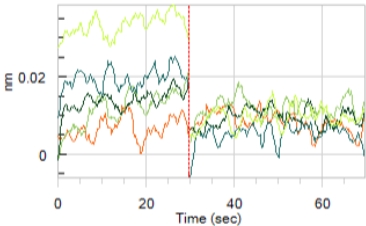
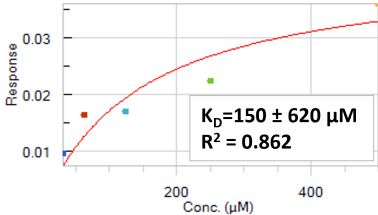
Continued on next page

Table 4.4 Continued from previous page

Hit	Structure	Binding	Steady State
2			
3			
4			
5			
6			

Continued on next page

Table 4.4 Continued from previous page

Hit	Structure	Binding	Steady State
7			
8			
9			
10			

Concluded



## 4.3 Discussion

A BLI binding assay was successfully established using biotinylated 6HisAvi*PaPanK*. BLI sensors could be reliably loaded with biotinylated 6HisAvi*PaPanK* and used to assess the binding of known ligands as well as a large number of compounds from a virtual screen and two fragment libraries. Based on analysis of the results, several hits were identified.

### 4.3.1 Expression and purification of 6HisAvi*PaPanK*

Biotinylation of *PaPanK* was necessary in order to load it to streptavidin-coated biosensors for BLI measurements. Site-specific biotinylation was desired to ensure that sensors could be loaded uniformly. Therefore, 6HisAvi*PaPanK*, which possesses a 6His-tag for purification by IMAC as well as an AviTag for site specific biotinylation was expressed [61].

Following initial purification of lysate utilizing a His-trap IMAC column, multiple peaks were observed (A.1). The low levels of imidazole (20 mM) present early in the purification seemed sufficient to prevent nonspecific binding to the column, and a vast majority of the cell lysate flowed through (peak A). As the concentration of imidazole increased (from 20 mM during sample application and washing, followed by gradient elution eventually reaching 500 mM imidazole), sample which had bound the column appeared to have been displaced for collection. High, > 80 mM concentrations of imidazole have been shown to displace 6His-tagged proteins by chelation of  $\text{Ni}^{2+}$ , and thus 6HisAvi*PaPanK* was suspected to be present in Peak B [62]. Later size-exclusion chromatography was used to further purify the protein and Peak C (Figure A.2) was suspected to contain the 6HisAvi*PaPanK*. Two smaller peaks, A and B, may have consisted of either denatured 6HisAvi*PaPanK*, 6HisAvi*PaPanK* which hydrolyzed at some point following His-trap purification, or non-desired proteins or compounds which bound to the His-trap column more strongly than expected. The SDS-PAGE performed on this suspected 6HisAvi*PaPanK* revealed a mostly pure band between the 25 and 37 kDa ladder bands (Figure 4.1), consistent with the 31.4 kDa mass of the recombinant protein. A fainter band was observed closer to the 25 kDa ladder bands and may indicate partial hydrolysis of the protein, resulting in a fraction which traversed the

polyacrylamide matrix slightly faster. It could also be due to a protein having co-eluted during IMAC and being too similar in size to be separated by SEC. Still, this protein was deemed adequately purified given that enzymatic biotinylation would introduce additional purification steps. Assuming efficient enzymatic biotinylation, the overall yield of 14.93 mg was greatly in excess of that required for many BLI experiments.

The biotinylation of 6HisAvi*Pa*PanK by BirA would later be confirmed by loading of the protein to SSA sensors.

### 4.3.2 BLI with *Pa*PanK

### 4.3.3 Protein loading optimization

Adsorption of biotinylated 6HisAvi*Pa*PanK onto SSA sensors was necessary prior to the assessment of small molecule binding. Given that BLI is designed to measure interference based on the optical density of bound molecules, a high degree of protein binding was desired in order to measure the likely low contribution to optical density provided by low molecular weight screening compounds. The biotinylated 6HisAvi*Pa*PanK loading took place over 5 minutes (Figure 4.2) to accommodate the diffusion limited  $K_{on}$  of biotin, reaching between 5 nm response at 12.5  $\mu\text{g}/\text{mL}$  and 11 nm at 250  $\mu\text{g}/\text{mL}$ . The dissociation step following protein loading showed only a slight drop in response after protein loading, indicating a minimal change in optical density and confirming that protein loading and enzymatic biotinylation of 6HisAvi*Pa*PanK were successful. The extremely low  $K_{off}$  of the biotin-streptavidin interaction [63] means that detachment from the sensor should be slow, though reference sensors are still needed to account for drift.

Concentrations of both 25  $\mu\text{g}/\text{mL}$  (7 nm response) and 100  $\mu\text{g}/\text{mL}$  (9.5 nm response) were considered to have adequate amounts of protein adsorption based on the work of Martin et al. (2011) where a protein loading response of 10 nm was reached after approximately 10 minutes for use in fragment screening [48]. While 250  $\mu\text{g}/\text{mL}$  did reach the highest levels of binding response observed, it required 10-fold more protein for a 1.6-fold increase in binding relative to loading at 25  $\mu\text{g}/\text{mL}$ . Thus, 25-100  $\mu\text{g}/\text{mL}$  biotinylated 6HisAvi*Pa*PanK was used for fragment screening and ligand analysis.

### Establishment of ATP and pantothenate binding

In order to use BLI for screening of fragment libraries and the virtual screening compounds, confirmation of the binding of confirmed *PaPanK* ligands was required. The binding of pantothenate was consistent with previous data, with a  $K_D$  of  $2.5 \pm 0.15 \mu\text{M}$  with two replicates (Figure 4.3d). Clear saturation was observed, with only minimal increases in response measured with successive doubling of concentration above  $25 \mu\text{M}$  (Figure 4.3, c). Although this was a low degree of confidence given only 2 measurements, similar values were observed with the use of pantothenate as a control throughout the fragment screening process. While 0.01% (v/v) Triton X-100 was included in follow-up experiments in order to limit aggregation, ATP measurements were not repeated with the inclusion of ATP and thus aggregation contributing to the high response cannot be ruled out. ATP is considered to be freely soluble in water [64], so aggregation may not be at play if the 5% (v/v) DMSO is compatible with ATP. Based on these results, pantothenate was considered validated for use as a positive control during fragment screening. ATP likely requires further validation despite its similarity to historical  $K_M$  values, as it was subject to limited testing.

### Fragment screen of virtual hits

Fragment screening of virtual hits attained by Lee (2020) revealed 10 compounds out of 17 which showed double-referenced BLI responses which exceeded 0.05 nm at at one or both of pH 6.0 and pH 7.5 (Table 4.1). This initial hit rate following preliminary screening is much higher than would be expected from a library-based fragment screen, but is not completely out of the realm of possibility as structure-based virtual screens have been performed with high rates of inhibitor identification [65].  $K_D$  determination of these 10 initial hits was carried out at both pH 6.0 and pH 7.5 and narrowed the field considerably, with only 2 compounds (Table 4.2), A1 and B1, being considered for further evaluation based on their responses and equilibrium dissociation constants at pH 7.5. Both showed 3-400  $\mu\text{M}$   $K_D$  values and individual concentration points fit well to the steady state plots (Figure 4.4). Maximum response values were lower when comparing the hit validation results (4.2) with the 2-point fragment screening results (4.1), which could be because a 4-fold lower concentration of  $\mu\text{g/mL}$  biotinylated 6HisAvi*PaPanK* was used in the hit validation experiment. The pantothenate positive

control responses were also lower given the reduced amount of protein loading, but still arrived at  $K_D$  values between 1.6 and 3.1  $\mu\text{M}$  with  $R_2 > 0.97$  across all tests (not shown).

The competitive inhibition experiment where pantothenate was used to outcompete any pre-bound compound A1 or B1 present in the *PaPanK* binding pocket gave somewhat ambiguous results (Figure 4.5). The binding responses of both A1 and pantothenate appeared to be fully additive, which indicated that they occupied entirely different binding pockets and likely ruled out compound A1 as a potential inhibitor. If they bound exclusively to the same pocket, the response was expected to be reduced when compared to at least one of the pantothenate-only or A1 only control association steps, yet this was not observed.  $K_D$  determination of pantothenate in the presence of constant 1250  $\mu\text{M}$  compound A1 seemed to confirm this, as the calculated  $K_D$  was lower in the presence of compound A1 than with pantothenate alone. In the case of compound B1 the response of pantothenate and B1 were only partially additive, perhaps indicating that B1 binds both to the binding pocket as well as elsewhere to the *PaPanK*. More B1 than was available would have been necessary to determine the  $K_D$  of pantothenate in the presence of B1, though such an experiment would have been highly informative. While not ideal as B1 appeared to compete at least partially with pantothenate, B1 was prioritized for crystallization experiments to determine its binding mode via x-ray crystallography.

### Fragment screen of EU-OS library

A subset of the EU-OS library was utilized in order to effectively test a much larger library, as the 1047 compounds were curated by the manufacturer to be representative of their  $> 100000$  compound fragment library. A more focused screen could then be conducted based on compounds in their repository with structural similarities to any confirmed hits. Despite single-concentration screening and hit-validation testing of these compounds, ultimately no compounds appeared promising enough for further analysis (Table 4.3). The presence of a carboxylate group as in pantothenate may be necessary for specific binding to the active site of *PaPanK*. The EU-OS contained only 2 carboxylic acids in total, and this may suggest that compounds lacking a carboxylate group (as is present in pantothenate) are unlikely to bind, and conversely that a carboxylate group may be necessary for strong binding to the active site.

### Fragment screen of a carboxylic acid library

Based on the importance of the carboxylate group of pantothenate in its interaction with the active site of *Pa*PanK, and its high incidence rate in the virtual screen compounds, a library consisting of carboxylic acids was tested via BLI. During KNIME refinement based on biocytin control responses and *Pa*PanK responses, the mean biocytin response  $\pm 1$  STD ( $.004 \pm .018$  nm) was considerably lower than that from the EU-OS library. It could be true that the increased solubility attributed to the carboxylate group reduces aggregation during BLI, resulting in fewer nonspecific binding events. The carboxylic acid compounds were specifically filtered beforehand to include only compounds with calculated solubility ( $c\text{LogS}$ )  $\leq -1.5$ . Indeed, only 4/302 carboxylic acids tests showed biocytin responses above 0.1 nm or below -0.1 nm, the extreme acceptable bounds used for initial refinement. Owing to this apparent lack of nonspecific interactions, few compounds showed *Pa*PanK responses greater than the mean + 1STD ( $0.012 + 0.039$  nm), and only 9 compounds (2.98 %) were identified through the KNIME workflow as candidates for hit validation. This lack of overall non-specificity could mean that only the most energetically favorable of interactions between the carboxylate and active site residues is possible, limiting the number of hits but with an increased likelihood of hits having real active site affinity. After visual evaluation, 10 compounds were ultimately hit-validated through serial dilutions and determination of  $K_D$  (Table 4.4).

Following serial dilutions of these compounds for  $K_D$  value determination (Table 4.4), 6 appeared to show readily saturable binding with sub millimolar  $K_D$  values (1, 4, 5, 6, 7, and 8), which includes 2 of the 4 compounds which were included manually despite elimination from the KNIME workflow (Table 4.4 hits 7 and 8), owing to the appearance of their association/dissociation curves and only marginally excessive biocytin responses. These 2 compounds showed the highest response values and lowest,  $< 200$   $\mu\text{M}$   $K_D$  values of the 6 compounds deemed acceptable, which could indicate that the KNIME workflow was too stringent. The KNIME workflow could be tweaked to account for more factors than response values alone, as the BLI also measures  $K_{\text{on}}$  and  $K_{\text{off}}$  rates. Further analysis of these compounds in order to determine binding site specificity is necessary before any of them can be used as a scaffold for antibiotic design.



# Chapter 5

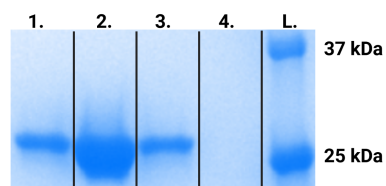
## *Pa*PanK protein crystallography

### 5.1 Expression and Purification of 6HisTEV*Pa*PanK

While BLI provides information about binding of compounds and can be used to narrow down a search for fragments with affinity for *Pa*PanK, it does not confirm active site specificity and does not provide binding mode information. Therefore the diffraction of *Pa*PanK crystals bound to fragments was desired. In order to produce *Pa*PanK crystals, a protein with a small, or removable purification tag was desired so as to not interfere with crystal formation. 6HisTEV*Pa*PanK was used, containing only a 6-peptide histidine purification tag and a TEV protease cleavage site in addition to the full *Pa*PanK primary sequence.

. In order to attempt crystallization of non-cleaved 6HisTEV*Pa*PanK, expression and purification were performed in the same manner as for 6HisAvi*Pa*PanK. Following purification by both a His-Trap and SEC, 16 mg of suspected 6HisTEV*Pa*PanK was recovered based on the ÄKTA's built-in UV meter and an  $A_{280}$  0.1% of 0.677 for a total yield of 4 mg purified protein per liter of *E. coli*. 3 adjacent elution fractions were run on a gel prior to pooling, and their sizes were consistent with the 29.4 kDa of 6HisTEV*Pa*PanK (Figure 5.1). Overloading appeared evident in lane 2 as the loaded amounts were not normalized.

In order to obtain *Pa*PanK without an affinity tag for use in additional crystallization trials 6HisTEV*Pa*PanK was expressed again, but was not subject to purification by



**Fig. 5.1:** SDS-PAGE for size and purity determination of 6HisTEV *PaPanK* following both His-Trap purification and SEC. Lanes 1-3 contain eluted SEC fractions suspected to contain 6HisTEV *PaPanK*. Lane 4 is empty. Lane L contains Precision Plus Protein™ Dual Xtra standard.

SEC. Following His-Trap purification, 39.45 mg of suspected 6HisTEV *PaPanK* was collected based on the built-in UV meter and an  $A_{280}$  0.1% of 0.677. Approximately half of this protein was later subject to tag-cleavage by TEV protease, and the protein prior to cleavage is shown in lane 4 in the following section (Figure 5.2)

### 5.1.1 TEV protease cleavage of 6HisTEV *PaPanK*

As an alternative to crystallizing *PaPanK* with a 6His-tag, TEV protease was used to cleave the added residues to allow for structural determination of the wild-type protein. The observation of large tag-cleaved *PaPanK* crystals in unpublished data provided further justification for tag cleavage. 6HisTEV *PaPanK* was treated with TEV protease at a protease : *PaPanK* protein mass ratio of 1 : 10 according to Method 3.7 at a temperature of 8C overnight. SDS-PAGE of the cleaved protein was performed prior to purification, and a band consistent with cleavage of the short 6His-tag was observed in Lane 3, when compared to the non-cleaved control in Lane 4 (5.2).



**Fig. 5.2:** SDS-PAGE for determination of 6His-tag cleavage stained with PageBlue. L) Protein ladder. 1) empty. 2) TEV protease normalized to lane 3. 3) *PaPanK* following TEV protease treatment. 4) 6HisTEV *PaPanK* prior to cleavage

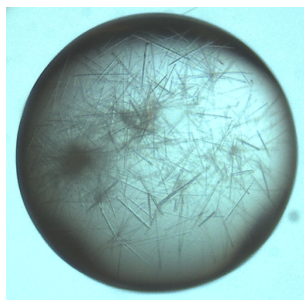
Individually cleaved sample tubes were subject to reverse His-trap purification as previously described (Method 3.7), and a large, isolated peak (Figure A.3 peak A) was observed during the washing phase with loading buffer (20 mM Tris HCl pH 7.9, 500 mM NaCl, 20mM Imidazole, 10% v/v glycerol). Based on the built in UV meter using the  $A_{280}$  0.1% of 0.677 (the value of which does not change upon cleavage), 12.47 mg



was recovered of an initial 17 mg input leading to a combined cleavage and recovery rate of 73%. SDS-PAGE was not performed on this protein following TEV cleavage and reverse His-Trap purification, but a small scale test (not shown) showed only an isolated, pure band which appeared to be smaller than the non-cleaved control.

## 5.2 Crystallization of *PaPanK*

Crystallization of *PaPanK* was desired in order to perform x-ray diffraction of *PaPanK* bound to ligands, including hits from BLI screening. A variety of commercial screening plates were tested on non-cleaved 6HisTEV*PaPanK* as described in Method 3.11, however these yielded only small needles that could pose difficulties for diffraction (Figure 5.3). Failed attempts were also made to crystallize 6HisTEV*PaPanK* in hanging drops per the method outlined in Hong et al. (2006), though the literature attempt differed in that it included Se-Met *PaPanK* with no affinity tag [35]. Thus, the crystallization of TEV-cleaved *PaPanK* was prioritized. A previous master's student expressed, purified, and cleaved 6HisTev*PaPanK* with TEV protease and observed the formation of needle-like crystals which did not diffract [49]. In a review of their unpublished data, it was discovered that they had produced larger, ~50  $\mu\text{M}$  wide singly occurring crystals using well E12 of the JCSG-plus crystallization buffer screen (100 mM Imidazole pH 8.0, 10% (w/v) PEG8000) and with *PaPanK* buffer exchanged to 50 mM Tris-HCl pH 7.5. Therefore, it was attempted to reproduce crystals under these conditions.



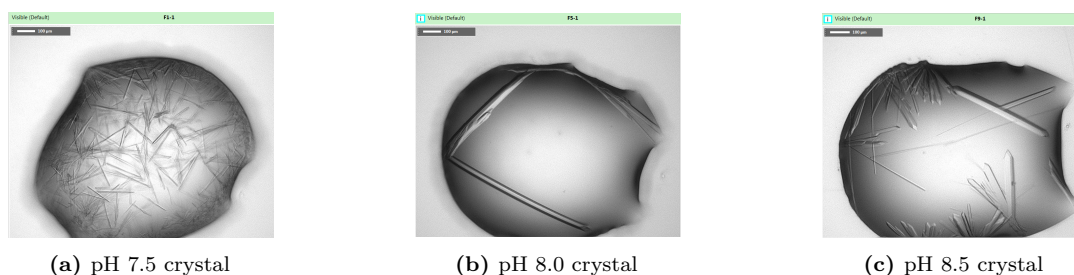
**Fig. 5.3:** A representative example of the numerous needle-like crystals which often formed when attempting to crystallize 6HisTEV*PaPanK* without cleavage of the 6His-tag. This particular crystal utilized 100 mM Bis-Tris Propane pH 6.5, 200 mM NaI, 10% (w/v) PEG3350, 10% (w/v) ethylene glycol, and 6HisTEV*PaPanK* at 22 mg/mL. A drop ratio of 150 nL protein:150 nL precipitant was used, and the crystal formed at 8 °C

In this work, frozen aliquots of this *PaPanK* were thawed on ice and buffer exchanged to 50 mM Tris-HCl pH 7.5 in order to remove glycerol. An optimization plate was produced, with varying protein concentrations, PEG concentrations, and precipitant pH values (Figure 5.4). Each of these 96 conditions varied also by protein : precipitant volume ratios, as either 100:200 nL, 150:150 nL, or 200:100 nL on the crystallization plate. Nucleation was almost immediate under many conditions, with visible crystals forming within the several minutes between sitting-drop placement and initial imaging. After 18 hours crystals at pH 7.5 were predominantly needles with many clustered crystals (Figure 5.5a). The lowest protein concentrations also appeared to favor needle formation under all conditions tested (full data not shown).

	1	2	3	4	5	6	7	8	9	10	11	12	mg/mL P
A													3.95
B					1				1				5.93
C					1	1							8.89
D					1, 3								11.85
E					1, 3		2						14.81
F					1			1	1				16.79
G					1, 3								18.76
H													19.75
	8%	9%	10%	11%	8%	9%	10%	11%	8%	9%	10%	11%	
	pH 7.5 imidazole 100mM				pH 8.0 imidazole 100mM				pH 8.5 imidazole 100mM				

**Fig. 5.4:** Plate layout of tag-cleaved *PaPanK* crystallization optimization. Each condition was tested at 3 different drop ratios as described, and crystal conditions selected for diffraction are highlighted in green with the drop ratio number included: '1' = 150 nL protein:150 nL precipitant, '2' = 100 nL protein:200 nL precipitant, '3' = 200 nL protein:100 nL precipitant. 'Mg/mL P' represents the protein concentration of *PaPanK* placed into drops. The varying percentage values represent the % (w/v) of PEG8000 in the precipitant.

The crystals which formed at pH 8.0 were generally the largest and least clustered. When comparing drop ratios, ratios of 150 nL protein : 150 nL precipitant resulted in similar crystals to those formed at a ratio of 200 : 100 (Figure 5.5b, c), however the

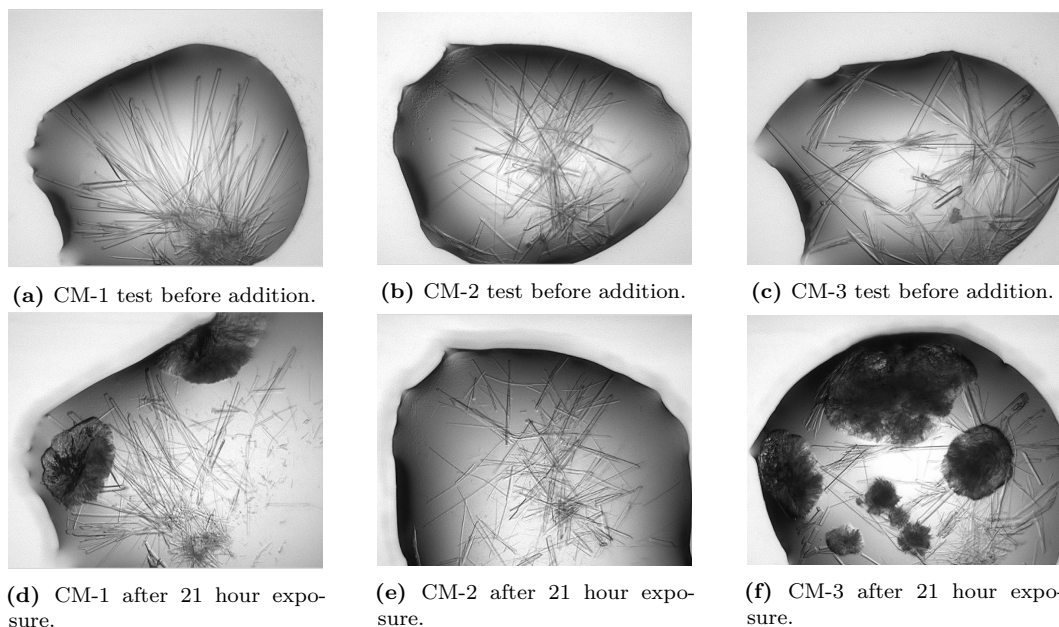


**Fig. 5.5:** Three His-tag cleaved *PaPanK* crystals imaged after 1 day 12 hours with 150 nL protein : 150 nL precipitant, 16.8 mg/mL protein, 100 mM imidazole, 8% (w/v) PEG8000 varying only by pH

latter crystals appeared to degrade more rapidly. Thus, condition F5-1 (Figure 5.5 b) was selected for use in future crystallography experiments.

### 5.2.1 Determination of suitable cryoprotectant

As described in Introduction 1.7, modern X-ray beamlines operate at temperatures below 120 K to reduce radiation damage, which makes cryoprotection necessary in order to avoid the formation of ice crystals. Although PEG8000 can act as a cryoprotectant, the previously chosen crystallization condition contained too little to be considered cryoprotected. Several cryoprotectants from the Cryosol MD1-90 (Molecular Dimensions) kit were added to *PaPanK* crystals as in Method 3.11.1. Only cryomixes (CM)-1, 2, and 3 were tested on intact crystals since the other crystal drops evaporated before the addition of cryomix. Images were taken of the crystals before and 21 hours after the addition of cryomixes, and the cryomixes were tolerated to varying degrees (Figure 5.6). Crystals appeared to have dissolved slightly under all conditions, while dark aggregates formed in CM-1 and CM-3. Each of the 3 cryomixes were selected for use as cryoprotectants in the proceeding trial in case one or more proved to hinder diffraction.



**Fig. 5.6:** Effect of Cryomix 1-3 on *PaPanK* crystals before treatment and after 21 hours of soaking

## 5.2.2 Diffraction and structural confirmation of *PaPanK* and various ligands

While condition F5-1 from the optimization plate (100 mM Imidazole pH 8.0, 8% w/v PEG8000, 150 nL protein : 150 nL well buffer, input *PaPanK* concentration 16.8 mg/mL) appeared most visually appealing, a number of other crystals from this plate also appeared suitable for attempting diffraction. After growing crystals, they were fished using loops and cryoprotected prior to flash-freezing and shipping to various beamline facilities. For the diffraction of apo-enzymes, crystals were cryoprotected by directly adding 500 nL of the selected cryomix solution diluted using the crystal-specific well buffer condition to the drop. Each of the 3 cryomix solutions which were tested on *PaPanK* crystals (Figure 5.6) were included with large crystals which formed in the optimization plate (Figure 5.4).

For several crystals, ligands at concentrations far exceeding experimental  $K_D$  values were included in the cryomix solutions and soaked by direct addition to the crystal drops for approximately 15 minutes (Figure 5.1). When soaking, the concentrations were 7.8 mM (A1), 7.8 mM (B1) and 6.5 mM (Pantothenate) within the crystal drops.

Flash-frozen protein crystals were beamed at the BioMAX MAX IV synchrotron in Lund, Sweden, and the resolutions at which crystals diffracted, along with ice-presence, are shown in Table 5.1. The majority of crystals which did not form ice diffracted at useful resolutions  $< 3.0$  Å, with 4/12 (33.34%) of crystals from which data was collected diffracting better than 2.0 Å. 6/7 crystals which utilized cryomix 1 for cryoprotection developed ice, and the two crystals which did diffract did under these conditions so at the lowest resolutions of 2.9 and 2.92 Å. Cryomix 1 was therefore excluded from further use as a cryoprotectant.

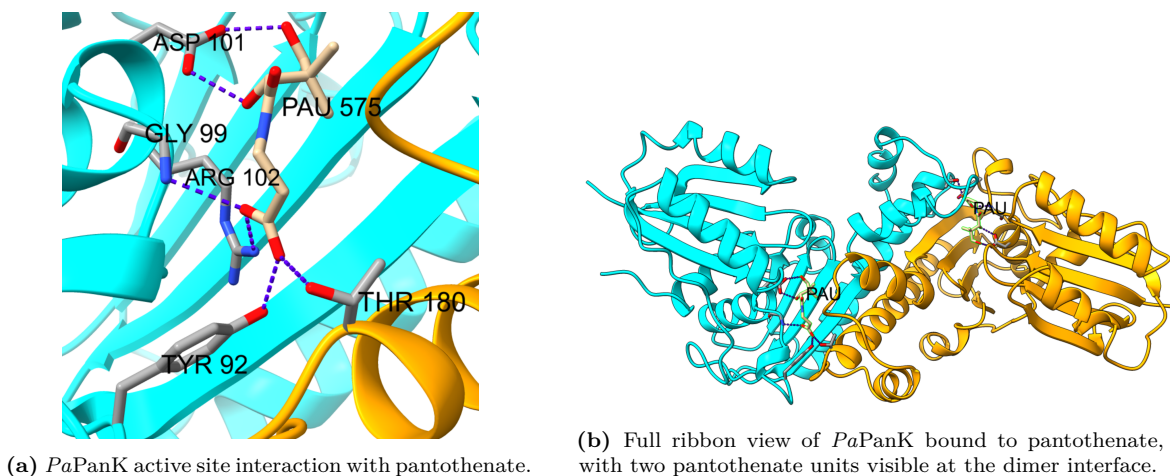
The collected data were further processed according to Method 3.11.4 and diffraction and refinement data are displayed in Table A.1. Only the structure of the *PaPanK* dimer was visible for the crystals which were soaked with virtual screen ligands A1 or B1, which indicated a lack of binding under these conditions. Complete structures of the apoenzyme as well as *PaPanK*-Pantothenate were determined (Figure 5.7). Pantothenate could be seen binding to both chains of the homodimer at the dimer interface with the same binding mode as previously shown in PDB 2F9W (Figure 5.7a) [35]. The carboxylate group of pantothenate formed hydrogen bonds with Tyr92 Gly99,

Crystal location	Ligand	Resolution (Å)/Ice	Cryomix
B5-1	Apo	Ice	1
B9-1	Apo	Ice	3
C5-1	Apo	Ice	1
C5-1	Apo	Ice	1
C6-1	Apo	1.9	2
C6-1	Apo	2.4	2
F5-1	Apo	2.92	1
F5-1	Apo	Ice	1
F8-1	Apo	2.9 /Ice	1
F8-1	Apo	Ice	1
F9-1	Apo	2.11	3
F9-1	Apo	No crystal	3
E7-2	Apo	1.9	3
E7-2	Apo	No crystal	3
D5-1	B1	2.23	3
D5-3	B1	2.27	3
D5-1	B1	Ice	3
E5-1	Pantothenate	1,95	3
E5-1	Pantothenate	1.82	3
E5-3	A1	2.8	3
G5-1	A1	2.03	3
G5-3	Pantothenate	2.11	3

**Table 5.1:** Data collection from crystals fished and optionally soaked with ligands from the optimization plate. Samples were soaked with varying cryomix solutions, while samples in blue were soaked with cryomix containing the highlighted ligand. B1 and A1 are hits from fragment screening of the virtual screen compounds (Section 4.2.3)

and Arg102 of one chain and Thr 180 of the opposing chain (Figure 5.7b). Asp101 stabilized the opposite end of pantothenate, interacting with two separate hydroxyl groups.

Various structural alignments were performed in order to gauge the level of similarity between historical structures and those obtained in this work (Table 5.2). The main chain residues of the apoenzyme and pantothenate-*PaPanK* were aligned with an rmsd of 0.642 Å. Alignment of the *PaPanK*-pantothenate with the previously published Se-Met *PaPanK*-Pantothenate (PDB 2F9w) gave an rmsd of 0.943 Å. The apoenzyme *PaPanK* did not show the locations of residues between Arg154 and Tyr162 on chain A, varying slightly from the apoenzyme Se-Met *PaPanK* (PDB 2F9T) which lacked



**Fig. 5.7:** (above) *PaPanK* homodimer bound to pantothenate, zoomed in to pantothenate binding site (a), and full view (b). PAU represents pantothenate, active site residues contributing hydrogen bonds (dashed lines) are labelled and displayed as sticks. Chain A is colored cyan and chain B is orange. Only *PaPanK* residues and pantothenate are displayed for clarity.

Reference structure	Alignment structure	rmsd (Å)
PaPanK-Pant	Apo PaPanK	0.642
Apo PaPanK Chain A	Apo PaPanK Chain B	0.718
PaPanK-Pant Chain A	PaPanK-Pant Chain B	0.824
PaPanK-Pant	Se-Met PaPanK-Pant (PDB 2F9W)	0.943
Apo PaPanK	Apo Se-Met PaPanK (PDB 2F9T)	1.066

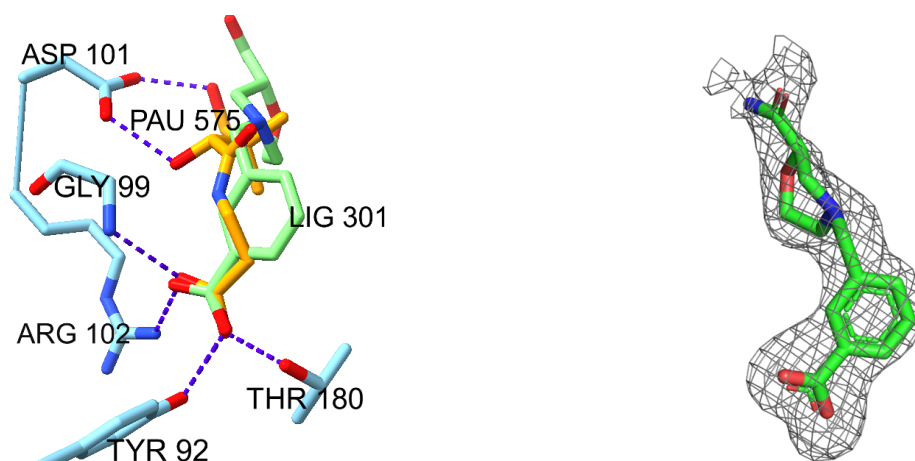
**Table 5.2:** rmsd (Å) values calculated in Chimera X using the Needleman-Wunsch sequence alignment algorithm for the *PaPanK* apoenzyme, *PaPanK*-pantothenate, Chain A and Chain B comparisons from within the same structures, and comparison with historical PDB entries which contained selenium-methionine in lieu of methionine. 'Apo' represents unbound apoenzymes while 'Pant' represents the ligand pantothenate. The best aligning chains were selected for comparison of the Se-Met and wild-type *PaPanK* structures.

information for the residues between Arg154 and Asp164. This non-mapped region was however present in the *PaPanK*-pantothenate and Se-Met *PaPanK*-pantothenate (PDB 2F9W) structures. Overall, there tended to be little variation between the aligned structures.



### Structural determination of *PaPanK*-ATP and *PaPanK*-B1

The diffraction of *PaPanK* crystals showing the binding modes of previously unpublished ligands was desired, as such information could be used to identify strongly binding ligands in the future. Repeated attempts were made to diffract *PaPanK* soaked with compound B1 from the virtual screen as well as a variety of other ligands. Soaking with pantothenate and AMP-PNP, an inactive ATP analogue, did not indicate electron density for AMP-PNP. A combination of ATP and pantothenate could not be soaked due to their reactivity. After several unsuccessful soaking and diffraction trials, evidence of compound B1 as well as ATP binding were seen in separate structures after performing an alternative soaking protocol per Method 3.11.3, in which the KCl concentration was increased and crystals were soaked in the absence of cryoprotectant. Data collections took place with the DESY P11 synchrotron in Hamburg, Germany. The data collection, processing and refinement statistics for *PaPanK*-B1 and *PaPanK*ATP are summarized in Appendix A.1.



(a) Aligned molecular models of diffracted *PaPanK* crystals obtained separately from virtual screen compound B1 and pantothenate. Only protein residues and H-bonds from the *PaPanK*-pantothenate structure are shown. PAU represents pantothenate while LIG represents compound B1.

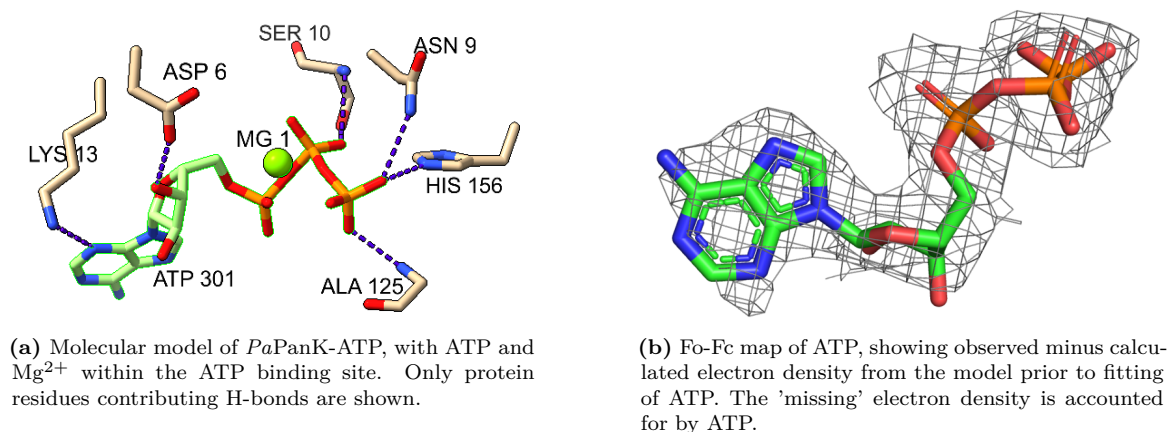
(b) Fo-Fc map of compound B1, with electron density with B1 excluded from the model shown as a grid.

**Fig. 5.8:** Preliminary x-ray structures for determination of the binding mode of compound B1. (a) Alignment of the previously shown *PaPanK*-Pantothenate structure with *PaPanK*-B1. (b) The Fo-Fc omit map for compound B1 used to show electron density without molecular model bias.

Preliminary refinement of the diffraction data of *PaPanK*-B1 revealed binding of compound B1 to the active site of *PaPanK*. Notably, the ligand was only visible at the dimer interface consisting mostly of chain A protein residues. An alignment of the

*PaPanK*-Pantothenate and *PaPanK*-B1 molecular models shows that both compounds occupy similar spaces within the active site, particularly where the carboxylate moieties are placed (Figure 5.8a). The preliminary model showed a relatively high ligand B-factor of  $81.04 \text{ \AA}^2$ , indicating a degree of uncertainty in the placement of B1. An Fo-Fc omit map was produced in order to determine whether electron density existed to justify the placement of B1 within the molecular model (Figure 5.8b). The Fo-Fc map mostly overlapped with the model of B1 within *PaPanK*, especially so at the carboxylate moiety. Electron density at the opposite end was less well-defined. Based on these observations, the approximate binding mode of virtual screening compound B1 appeared to be confirmed.

Preliminary refinement of the diffraction data of *PaPanK*-ATP revealed binding of ATP, which like B1 was visible at only one of two binding sites. Superimposition of the *PaPanK*-ATP and *PaPanK*-Pantothenate chain A structures had an rmsd of  $0.483 \text{ \AA}$ , indicating a high degree of similarity between both structures. The predicted molecular model also showed a relatively high B-factor of (Table A.1).



**Fig. 5.9:** Preliminary x-ray structures for determination of the binding mode of compound B1. (a) Alignment of the previously shown *PaPanK*-Pantothenate structure with *PaPanK*-B1. (b) The Fo-Fc omit map for ATP used to show electron density with reduced molecular model bias. The molecule has been rotated relative to figure (a) in order to more clearly display the electron density.



## 5.3 Discussion: Protein crystallography

Well-diffracting *PaPanK* crystals were desirable as these could be used to assess the binding modes of hits which resulted from BLI, as well as natural ligands which have not yet been structurally determined when bound to *PaPanK*. 6HisTev*PaPanK* was expressed and purified as it was hoped that its small tag would not heavily interfere with crystallization, and could be removed by TEV protease if necessary. Ultimately a number of large *PaPanK* crystals were produced, which were successfully used for structural determination of the binding modes of various ligands at  $\sim 2.0$  Å or better, and at times served as an excellent complement to the BLI measurements.

### 5.3.1 Expression and purification of 6HisTEV*PaPanK*

In order to collect structural x-ray data from *PaPanK* and various ligands, the expression of the protein was first required so that crystallization attempts could be made. The protein was expressed without issue, with no optimization required as the protocol had been previously established [49]. TEV protease cleavage of the 6His-tag was also considered successful, with a relatively pure 27 kDa band (Figure 5.2, lane 3) prior to reverse His-trap purification. While reverse His-trap revealed a singular peak in the fraction expected to contain protein without a His-tag (Figure A.3, Peak A), SDS-PAGE should have been performed for final size and purity confirmation prior to crystallization attempts. The identity of this purified protein was later confirmed to be tag-cleaved *PaPanK* by x-ray diffraction.

### 5.3.2 Crystallization of *PaPanK*

The formation of crystals of either *PaPanK* or 6HisTEV*PaPanK* was desired in order to facilitate x-ray diffraction. Attempts to crystallize 6HisTEV*PaPanK* yielded at best small needles, which would have posed difficulties both for fishing and diffraction. It has previously been shown that N-terminal 6His-tags can be subject to variable degrees of post-translational modifications in *E. coli* that can interfere with crystal growth [66]. This could be a possibility, or it could be that by chance none of the conditions tested were ideally suited for the formation of large crystals. When utilizing

tag-cleaved *PaPanK* and variations based on well E12 of the JCSG-plus screen, large crystals were readily producible. The absence of a 6His-tag was more desirable as only the amino acid residues expressed by the *CoaX* gene would be present in any successfully diffracted crystals. Cryoprotectant testing was conducted in order to prevent ice crystals which might otherwise interfere with diffraction. The 21 hour test revealed that the *PaPanK* crystals could tolerate each of the 3 applied cryoprotectants (Figure 5.6). Dark aggregates appeared in CM-1 and CM-3 which both contained DMSO, but given that 21 hours far exceeded the planned ligand/cryoprotectant soaking time, all 3 were considered for later use.

### 5.3.3 Diffraction and structural confirmation of *PaPanK* and various ligands

#### Diffraction of *PaPanK* and *PaPanK*-Pantothenate

In order to develop a pipeline for the structural determination of ligand-*PaPanK* interactions, well-diffracting crystals were required. With the crystallization of *PaPanK* established, attempts were made to diffract the protein both as an apoenzyme and following soaking with cryomixes and ligands. A number of crystals indicated diffraction at useful resolutions (Table 5.1) prior to further processing and refinement. Following refinement of the molecular models generated from the electron density maps, molecular models were produced which aligned well with the historical Se-Met *PaPanK* and Se-Met *PaPanK*-pantothenate data, with fairly low rmsd (Å) (Figure 5.2). No major structural differences were apparent upon visual comparison of the molecular models with the historical Se-Met structures, and the disordered region beginning at Arg154 was observed in both apoenzymes but not the pantothenate-bound enzymes. Overall, the results from this trial confirmed that the methodology used in this work could produce well-diffracting *PaPanK* crystals.

#### Diffraction of virtual screening compounds

The first attempt at diffraction of the 2 most promising virtual screening compounds (A1 and B1) was unsuccessful in identifying any electron density that would implicate their

presence in the *PaPanK* structure, and only the apoenzymes were visible. Compound A1 was subsequently ruled out based upon its apparent lack of competitive inhibition with pantothenate and response-additivity during BLI (Figure 4.5). After several trials, preliminary data suggested electron density for compound B1 in the active site of *PaPanK*. As opposed to earlier soaking attempts, the successful attempt utilized a 56 mM KCl concentration during soaking and was soaked in the absence of cryoprotectant. The preliminary *PaPanK*-B1 data indicated a high ligand B-factor of 81.04 Å<sup>2</sup>, relative to the protein atoms B-factor of 58.63 Å<sup>2</sup>. This indicates that the certainty of the location of individual B1 atoms within the active site is considerably less than that of the surrounding protein atoms, and contrasts with the *PaPanK*-pantothenate structure in which the ligand B-factor was lower than the protein atom B-factor. This is not entirely unexpected given the weak binding indicated during BLI (B1 K<sub>D</sub>=410 μM), as weakly binding compounds are not expected to completely bind to all predicted binding sites and must compete for the binding sites with all other components used during crystallization [67].

While a K<sub>D</sub> value of 410 μM isn't especially weak when compared with ATP, it is possible that the affinity is weaker than calculated given the partially additive BLI response when combined with pantothenate as described in Discussion 4.3.3. Still, the binding of B1 appears to be confirmed based on the omit map modelled without B1, which clearly shows an area of electron density in the active site into which B1 can be readily accommodated (Figure 5.8b). Overlaying the *PaPanK*-model with the preliminary *PaPanK*-B1 model also shows that the carboxylate groups of these two ligands occupy very similar spaces within the active site. As evidenced by the binding mode of pantothenate, the conservation of this interaction in B1 underlines the importance of the carboxylate group in *PaPanK* active site affinity. Further, this result appears to validate the methodology used for soaking with relatively weak ligands and shows that the BLI screening methodology can identify small compounds with *PaPanK* active site specificity. The fact that B1 was accurately predicted the virtual screen conducted by Lee (2020) to bind to the active site of *PaPanK* coincides well with the low rmsd (Å) between the previous Se-Met *PaPanK* models and the wild-type models reported in this work (Table 5.2) [49]. Were the model highly inaccurate, simulated binding modes might be expected to be inaccurate as well.

### Diffraction of *Pa*PanK-ATP

Based on a preliminary refinement of diffraction data, ATP appeared to be present at one of two *Pa*PanK ATP binding sites (Figure 5.9a). The Fo-Fc omit map (Figure 5.9b) strongly indicates its presence, though still in need of further refinement. The current lack of placement of water molecules in the refined model (Appendix A.1) likely contributes to the relatively high  $R_{\text{free}}$  of 0.25 when compared to the more refined *Pa*PanK and *Pa*PanK-Pantothenate structures [68]. A high ligand B-factor relative to the protein atoms (78.58 Å<sup>2</sup> versus 58.63 Å<sup>2</sup>) indicates flexibility or uncertainty in the ligand placement due to reduced/dispersed electron density. The millimolar  $K_D$  and  $K_M$  values of ATP likely contribute to this high B-factor relative to the protein atoms, possibly due to low occupancy or the high degree of movement allowed for by the lack of strongly stabilizing interactions. Yang et al. (2008) also observed a considerably higher B-factor for ADP than for protein residues in the structure of *Thermotoga maritima* PanK-Pantothenate-ADP, another type III PanK with a similarly low affinity for ATP. Further, the observed binding of ATP in the *Pa*PanK-ATP structure is consistent with the observation by Yang et al. (2008) that *Tm*PanK bound AMP-PNP with equally low affinity regardless of the presence or absence of pantothenate [33]. As previously mentioned, the *Pa*PanK-ATP model requires further refinement, but its ability to bind *Pa*PanK in the absence of pantothenate appears to be confirmed.

# Chapter 6

## Conclusion

Overall, considerable progress was made based on the original objectives of this thesis. BLI utilizing enzymatically biotinylated 6HisAvi*PaPanK* was successfully implemented as a tool for screening large numbers of fragments with potential *PaPanK* specificity. *PaPanK* crystals which consistently diffract at useful resolutions were produced, and a method for soaking these *PaPanK* crystals with ligands as weakly binding as ATP was developed. The molecular models produced from these diffracted crystals included evidence for the binding mode of ATP in the absence of pantothenate, as well as the binding of a previous virtual screening compound, B1. The binding of this compound served not just to validate the virtual and fragment screening processes, but also provides a scaffold for chemical modifications which could be used to obtain a more strongly binding ligand in the future, or to curate a library for similar compounds. The 6 hits from the carboxylic acid library have yet to be soaked with *PaPanK* crystals, so there exists the potential for determining whether those compounds bind as well. Future work on this topic could even adapt the methods developed here to conduct high-throughput screening by X-ray crystallography.



# References

- [1] William D. Fiers, Mark Craighead, and Ishwar Singh. Teixobactin and Its Analogues: A New Hope in Antibiotic Discovery. *ACS Infectious Diseases*, 3(10):688–690, October 2017. Publisher: American Chemical Society.
- [2] World Health Organization et al. Who priority pathogens list for r&d of new antibiotics. *World Health Organization, Geneva, Switzerland*, 2017.
- [3] Jens Klockgether and Burkhard Tümmler. Recent advances in understanding *Pseudomonas aeruginosa* as a pathogen. *F1000Research*, 6, July 2017.
- [4] Cassiano Felipe Gonçalves-de Albuquerque, Adriana Ribeiro Silva, Patrícia Burth, Patricia Rieken Macêdo Rocco, Mauro Velho Castro-Faria, and Hugo Caire Castro-Faria-Neto. Possible mechanisms of *Pseudomonas aeruginosa*-associated lung disease. *International Journal of Medical Microbiology*, 306(1):20–28, January 2016.
- [5] S. P. Lilani, N. Jangale, A. Chowdhary, and G. B. Daver. Surgical site infection in clean and clean-contaminated cases. *Indian Journal of Medical Microbiology*, 23(4):249–252, October 2005.
- [6] Andrea D. Valderrey, María José Pozuelo, Pedro A. Jiménez, María D. Maciá, Antonio Oliver, and Rafael Rotger. Chronic colonization by *Pseudomonas aeruginosa* of patients with obstructive lung diseases: cystic fibrosis, bronchiectasis, and chronic obstructive pulmonary disease. *Diagnostic Microbiology and Infectious Disease*, 68(1):20–27, September 2010.
- [7] Tzyy-Bin Tsay, Yu-Zhen Jiang, Ching-Mei Hsu, and Lee-Wei Chen. *Pseudomonas aeruginosa* colonization enhances ventilator-associated pneumonia-induced lung injury. *Respiratory Research*, 17, 2016.
- [8] Oded Orgad, Yoram Oren, Sharon L. Walker, and Moshe Herzberg. The role of alginate in *Pseudomonas aeruginosa* EPS adherence, viscoelastic properties and cell attachment. *Biofouling*, 27(7):787–798, July 2011. Publisher: Taylor & Francis \_eprint: <https://doi.org/10.1080/08927014.2011.603145>.
- [9] Liang Yang, Martin Nilsson, Morten Gjermansen, Michael Givskov, and Tim Tolker-Nielsen. Pyoverdine and PQS mediated subpopulation interactions involved

- in *Pseudomonas aeruginosa* biofilm formation. *Molecular Microbiology*, 74(6):1380–1392, 2009. \_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2958.2009.06934.x>.
- [10] George A. O’Toole and Roberto Kolter. Flagellar and twitching motility are necessary for *Pseudomonas aeruginosa* biofilm development. *Molecular Microbiology*, 30(2):295–304, 1998. \_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1046/j.1365-2958.1998.01062.x>.
- [11] Victoria A. Marko, Sara L. N. Kilmury, Lesley T. MacNeil, and Lori L. Burrows. *Pseudomonas aeruginosa* type IV minor pilins and PilY1 regulate virulence by modulating FimS-AlgR activity. *PLOS Pathogens*, 14(5):e1007074, May 2018. Publisher: Public Library of Science.
- [12] Guy R. Cornelis and Frédérique Van Gijsegem. Assembly and Function of Type III Secretory Systems. *Annual Review of Microbiology*, 54(1):735–774, October 2000. Publisher: Annual Reviews.
- [13] Philipp Wolf and Ursula Elsässer-Beile. *Pseudomonas* exotoxin A: From virulence factor to anti-cancer agent. *International Journal of Medical Microbiology*, 299(3):161–176, March 2009.
- [14] Deepak Balasubramanian, Lisa Schneper, Massimo Merighi, Roger Smith, Giri Narasimhan, Stephen Lory, and Kalai Mathee. The Regulatory Repertoire of *Pseudomonas aeruginosa* AmpC  $\beta$ -Lactamase Regulator AmpR Includes Virulence Genes. *PLOS ONE*, 7(3):e34067, March 2012. Publisher: Public Library of Science.
- [15] Jerome Birnbaum, Frederick M. Kahan, Helmut Kropp, and James S. Macdonald. Carbapenems, a new class of beta-lactam antibiotics: Discovery and development of imipenem/cilastatin. *The American Journal of Medicine*, 78(6, Supplement 1):3–21, June 1985.
- [16] G Meletis, M Exindari, N Vavatsi, D Sofianou, and E Diza. Mechanisms responsible for the emergence of carbapenem resistance in *Pseudomonas aeruginosa*. *Hippokratia*, 16(4):303–307, 2012.
- [17] David M Livermore. Of pseudomonas, porins, pumps and carbapenems. *Journal of Antimicrobial Chemotherapy*, 47(3):247–250, 2001.
- [18] Barbara S. Margaret, G. L. Drusano, and Harold C. Standiford. Emergence of resistance to carbapenem antibiotics in *Pseudomonas aeruginosa*. *Journal of Antimicrobial Chemotherapy*, 24(suppl\_A):161–167, January 1989.
- [19] George G. Zhanel, Ryan Wiebe, Leanne Dilay, Kristjan Thomson, Ethan Rubinstein, Daryl J. Hoban, Ayman M. Noreddin, and James A. Karlowsky. Comparative Review of the Carbapenems. *Drugs*, 67(7):1027–1052, May 2007.
- [20] José-Manuel Rodríguez-Martínez, Laurent Poirel, and Patrice Nordmann. Molecular epidemiology and mechanisms of carbapenem resistance in *Pseudomonas aeruginosa*. *Antimicrobial agents and chemotherapy*, 53(11):4783–4788, 2009.



- [21] Alessandro Cassini, Liselotte Diaz Högberg, Diamantis Plachouras, Annalisa Quattrocchi, Ana Hoxha, Gunnar Skov Simonsen, Mélanie Colomb-Cotinot, Mirjam E Kretzschmar, Brecht Devleesschauwer, Michele Cecchini, Driss Ait Ouakrim, Tiago Cravo Oliveira, Marc J Struelens, Carl Suetens, Dominique L Monnet, Reinhild Strauss, Karl Mertens, Thomas Struyf, Boudewijn Catry, Katrien Latour, Ivan N Ivanov, Elina G Dobрева, Arjana Tambic Andrašević, Silvija Soprek, Ana Budimir, Niki Paphitou, Helena Žemlicková, Stefan Schytte Olsen, Ute Wolff Sönksen, Pille Märtin, Marina Ivanova, Outi Lyytikäinen, Jari Jalava, Bruno Coignard, Tim Eckmanns, Muna Abu Sin, Sebastian Haller, George L Daikos, Achilleas Gikas, Sotirios Tsiodras, Flora Kontopidou, Ákos Tóth, Ágnes Hajdu, Ólafur Guðlaugsson, Karl G Kristinsson, Stephen Murchan, Karen Burns, Patrizio Pezzotti, Carlo Gagliotti, Uga Dumpis, Agne Liuimiene, Monique Perrin, Michael A Borg, Sabine C de Greeff, Jos CM Monen, Mayke BG Koek, Petter Elstrøm, Dorota Zabicka, Aleksander Deptula, Waleria Hryniewicz, Manuela Caniça, Paulo Jorge Nogueira, Paulo André Fernandes, Vera Manageiro, Gabriel A Popescu, Roxana I Serban, Eva Schréterová, Slavka Litvová, Mária Štefkovicová, Jana Kolman, Irena Klavs, Aleš Korošec, Belén Aracil, Angel Asensio, María Pérez-Vázquez, Hanna Billström, Sofie Larsson, Jacqui S Reilly, Alan Johnson, and Susan Hopkins. Attributable deaths and disability-adjusted life-years caused by infections with antibiotic-resistant bacteria in the EU and the European Economic Area in 2015: a population-level modelling analysis. *The Lancet Infectious Diseases*, 19(1):56–66, January 2019.
- [22] Yu Zhang, Xiao-Li Chen, Ai-Wei Huang, Su-Ling Liu, Wei-Jiang Liu, Ni Zhang, and Xu-Zai Lu. Mortality attributable to carbapenem-resistant *Pseudomonas aeruginosa* bacteremia: a meta-analysis of cohort studies. *Emerging Microbes & Infections*, 5(1):1–6, January 2016. Publisher: Taylor & Francis \_eprint: <https://doi.org/10.1038/emi.2016.22>.
- [23] Burkhard Tümmler. Emerging therapies against infections with *Pseudomonas aeruginosa*. *F1000Research*, 8, August 2019.
- [24] Rehab M Abd El-Baky, Salwa M Masoud, Doaa S Mohamed, Nancy GFM Waly, Engy A Shafik, Dina A Mohareb, Azza Elkady, Mohamed M Elbadr, and Helal F Hetta. Prevalence and Some Possible Mechanisms of Colistin Resistance Among Multidrug-Resistant and Extensively Drug-Resistant *Pseudomonas aeruginosa*. *Infection and Drug Resistance*, 13:323–332, February 2020.
- [25] George Dimopoulos, Murat Akova, Jordi Rello, and Garyphalia Poulakou. Understanding resistance in *Pseudomonas*. *Intensive Care Medicine*, 46(2):350–352, February 2020.
- [26] María Díez-Aguilar, Marta Hernández-García, María-Isabel Morosini, Ad Fluit, Michael M Tunney, Natalia Huertas, Rosa del Campo, Daniel Obrecht, Francesca Bernardini, Miquel Ekkelenkamp, and Rafael Cantón. Murepavadin antimicrobial activity against and resistance development in cystic fibrosis *Pseudomonas aeruginosa* isolates. *Journal of Antimicrobial Chemotherapy*, 76(4):984–992, April 2021.

- [27] Tadhg P Begley, Cynthia Kinsland, and Erick Strauss. The biosynthesis of coenzyme a in bacteria. In *Vitamins & Hormones*, volume 61 of *Cofactor Biosynthesis*, pages 157–171. Academic Press, January 2001.
- [28] ROBERTA LEONARDI and SUZANNE JACKOWSKI. Biosynthesis of Pantothenic Acid and Coenzyme A. *EcoSal Plus*, 2(2), April 2007.
- [29] Joshua B Parsons and Charles O Rock. Is bacterial fatty acid synthesis a valid target for antibacterial drug discovery? *Current Opinion in Microbiology*, 14(5):544–549, October 2011.
- [30] Frederica L. Theodoulou, Ody C.M. Sibon, Suzanne Jackowski, and Ivan Gout. Coenzyme A and its derivatives: renaissance of a textbook classic. *Biochemical Society Transactions*, 42(4):1025–1032, August 2014.
- [31] Roberta Leonardi, Yong-Mei Zhang, Charles O. Rock, and Suzanne Jackowski. Coenzyme A: Back in action. *Progress in Lipid Research*, 44(2):125–153, March 2005.
- [32] Sara Cheek, Krzysztof Ginalski, Hong Zhang, and Nick V. Grishin. A comprehensive update of the sequence and structure classification of kinases. *BMC Structural Biology*, 5(1):6, March 2005.
- [33] Kun Yang, Erick Strauss, Carlos Huerta, and Hong Zhang. Structural Basis for Substrate Binding and the Catalytic Mechanism of Type III Pantothenate Kinase. *Biochemistry*, 47(5):1369–1380, February 2008. Publisher: American Chemical Society.
- [34] Leisl A. Brand and Erick Strauss. Characterization of a New Pantothenate Kinase Isoform from *Helicobacter pylori*\*. *Journal of Biological Chemistry*, 280(21):20185–20188, May 2005. Publisher: Elsevier.
- [35] Bum Soo Hong, Mi Kyung Yun, Yong-Mei Zhang, Shigeru Chohnan, Charles O. Rock, Stephen W. White, Suzanne Jackowski, Hee-Won Park, and Roberta Leonardi. Prokaryotic Type II and Type III Pantothenate Kinases: The Same Monomer Fold Creates Dimers with Distinct Catalytic Properties. *Structure*, 14(8):1251–1261, August 2006.
- [36] Madhav Mondhe, Ashley Chessher, Shan Goh, Liam Good, and James E. M. Stach. Species-Selective Killing of Bacteria by Antimicrobial Peptide-PNAs. *PLOS ONE*, 9(2):e89082, February 2014. Publisher: Public Library of Science.
- [37] Aurijit Sarkar and Ruth Brenk. To Hit or Not to Hit, That Is the Question – Genome-wide Structure-Based Druggability Predictions for *Pseudomonas aeruginosa* Proteins. *PLOS ONE*, 10(9):e0137279, September 2015. Publisher: Public Library of Science.
- [38] Keith H Turner, Aimee K Wessel, Gregory C Palmer, Justine L Murray, and Marvin Whiteley. Essential genome of *Pseudomonas aeruginosa* in cystic fibrosis sputum. *Proceedings of the National Academy of Sciences*, 112(13):4110–4115, 2015.

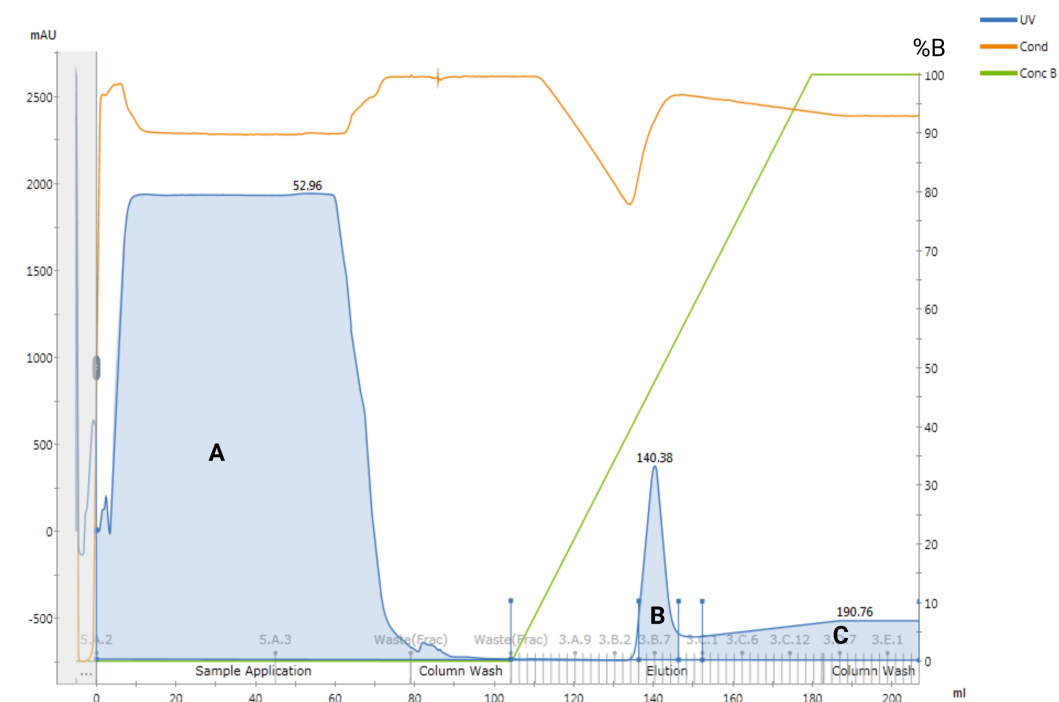
- [39] Samuel A Lee, Larry A Gallagher, Metawee Thongdee, Benjamin J Staudinger, Soyeon Lippman, Pradeep K Singh, and Colin Manoil. General and condition-specific essential functions of *Pseudomonas aeruginosa*. *Proceedings of the National Academy of Sciences*, 112(16):5189–5194, 2015.
- [40] Bradley E Poulsen, Rui Yang, Anne E Clatworthy, Tiantian White, Sarah J Osmulski, Li Li, Cristina Penaranda, Eric S Lander, Noam Shores, and Deborah T Hung. Defining the core essential genome of *Pseudomonas aeruginosa*. *Proceedings of the National Academy of Sciences*, 116(20):10072–10080, 2019.
- [41] Carl J Balibar, Micah F Hollis-Symynkywicz, and Jianshi Tao. Pantethine rescues phosphopantothencysteine synthetase and phosphopantothencysteine decarboxylase deficiency in escherichia coli but not in *Pseudomonas aeruginosa*. *Journal of bacteriology*, 193(13):3304–3312, 2011.
- [42] Jinming Guan, Leanne Barnard, Jeanne Cresson, Annabelle Hoegl, Justin H. Chang, Erick Strauss, and Karine Auclair. Probing the ligand preferences of the three types of bacterial pantothenate kinase. *Bioorganic & Medicinal Chemistry*, 26(22):5896–5902, December 2018.
- [43] Fredrik N. B. Edfeldt, Rutger H. A. Folmer, and Alexander L. Breeze. Fragment screening to predict druggability (ligandability) and lead discovery success. *Drug Discovery Today*, 16(7-8):284–287, April 2011.
- [44] Philine Kirsch, Alwin M. Hartman, Anna K. H. Hirsch, and Martin Empting. Concepts and Core Principles of Fragment-Based Drug Design. *Molecules*, 24(23):4309, January 2019. Number: 23 Publisher: Multidisciplinary Digital Publishing Institute.
- [45] Suzanne B. Shuker, Philip J. Hajduk, Robert P. Meadows, and Stephen W. Fesik. Discovering High-Affinity Ligands for Proteins: SAR by NMR. *Science*, 274(5292):1531–1534, November 1996. Publisher: American Association for the Advancement of Science Section: Reports.
- [46] Robin A. E. Carr, Miles Congreve, Christopher W. Murray, and David C. Rees. Fragment-based lead discovery: leads by design. *Drug Discovery Today*, 10(14):987–992, July 2005.
- [47] Martin Nirschl, Florian Reuter, and Janos Vörös. Review of Transducer Principles for Label-Free Biomolecular Interaction Analysis. *Biosensors*, 1(3):70–92, July 2011.
- [48] Eric Martin, John Wang, Isabel Zaror, Jiamin Yu, Kelly Yan, Mike Doyle, Paul Feucht, Kevin Shoemaker, Bob Warne, Mike Chin, Blisseth Sy, Lukas Leder, Marco Meyerhofer, Charles Wartchow, and Danfeng Yao. Novartis Evaluation of the ForteBio Octet RED: A Versatile Instrument for Direct Binding Experiments. In *Label-Free Technologies for Drug Discovery*, pages 223–240. John Wiley & Sons, Ltd. Section: 15 \_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/9780470979129.ch15>.

- [49] Ying Ho Leon Lee. Structure-based design of pantothenate kinase inhibitors as lead structures for new antibiotics. Master's thesis, The University of Bergen, 2020.
- [50] Alexander Wlodawer, Wlodek Minor, Zbigniew Dauter, and Mariusz Jaskolski. Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures. *The FEBS journal*, 275(1):1–21, January 2008.
- [51] Christopher G. Pope. X-Ray Diffraction and the Bragg Equation. *Journal of Chemical Education*, 74(1):129, January 1997. Publisher: American Chemical Society.
- [52] G. Taylor. The phase problem. *Acta Crystallographica Section D: Biological Crystallography*, 59(11):1881–1890, November 2003. Number: 11 Publisher: International Union of Crystallography.
- [53] Garry Taylor. The phase problem. *Acta Crystallographica Section D*, 59(11):1881–1890, Nov 2003.
- [54] Z. Dauter. Data-collection strategies. *Acta Crystallographica Section D: Biological Crystallography*, 55(10):1703–1717, October 1999. Number: 10 Publisher: International Union of Crystallography.
- [55] C. Gati, G. Bourenkov, M. Klinge, D. Rehders, F. Stellato, D. Oberthür, O. Yefanov, B. P. Sommer, S. Mogk, M. Duszynski, C. Betzel, T. R. Schneider, H. N. Chapman, and L. Redecke. Serial crystallography on in vivo grown microcrystals using synchrotron radiation. *IUCrJ*, 1(2):87–94, March 2014. Number: 2 Publisher: International Union of Crystallography.
- [56] D. Flot, T. Mairs, T. Giraud, M. Guijarro, M. Lesourd, V. Rey, D. van Brussel, C. Morawe, C. Borel, O. Hignette, J. Chavanne, D. Nurizzo, S. McSweeney, and E. Mitchell. The ID23-2 structural biology microfocus beamline at the ESRF. *Journal of Synchrotron Radiation*, 17(1):107–118, January 2010. Number: 1 Publisher: International Union of Crystallography.
- [57] Dimo Kashchiev. *Nucleation*. Elsevier, 2000.
- [58] Alexander McPherson and Jose A. Gavira. Introduction to protein crystallization. *Acta Crystallographica. Section F, Structural Biology Communications*, 70(Pt 1):2–20, December 2013.
- [59] Neer Asherie. Protein crystallization and phase diagrams. *Methods*, 34(3):266–272, November 2004.
- [60] U. K. Laemmli. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature*, 227(5259):680–685, August 1970. Number: 5259 Publisher: Nature Publishing Group.
- [61] Michael Fairhead and Mark Howarth. Site-specific biotinylation of purified proteins using BirA. *Methods in molecular biology (Clifton, N.J.)*, 1266:171–184, 2015.

- [62] R Janknecht, G de Martynoff, J Lou, R A Hipskind, A Nordheim, and H G Stunnenberg. Rapid and efficient purification of native histidine-tagged protein expressed by recombinant vaccinia virus. *Proceedings of the National Academy of Sciences*, 88(20):8972–8976, 1991.
- [63] O. H. Laitinen, V. P. Hytönen, H. R. Nordlund, and M. S. Kulomaa. Genetically engineered avidins and streptavidins. *Cellular and Molecular Life Sciences CMLS*, 63(24):2992–3017, December 2006.
- [64] PubChem. Adenosine-5'-triphosphate.
- [65] Zhe Wang, Huiyong Sun, Chao Shen, Xueping Hu, Junbo Gao, Dan Li, Dongsheng Cao, and Tingjun Hou. Combined strategies in structure-based virtual screening. *Physical Chemistry Chemical Physics*, 22(6):3149–3159, 2020. Publisher: Royal Society of Chemistry.
- [66] Kristine M Kim, Eugene C Yi, David Baker, and Kam YJ Zhang. Post-translational modification of the n-terminal his tag interferes with the crystallization of the wild-type and mutant sh3 domains from chicken src tyrosine kinase. *Acta Crystallographica Section D: Biological Crystallography*, 57(5):759–762, 2001.
- [67] Marc C Deller and Bernhard Rupp. Models of protein–ligand crystal structures: trust, but verify. *Journal of computer-aided molecular design*, 29(9):817–836, 2015.
- [68] Ivan G. Shabalin, Przemyslaw J. Porebski, and Wladek Minor. Refining the macromolecular model – achieving the best agreement with the data from X-ray diffraction experiment. *Crystallography reviews*, 24(4):236–262, 2018.



# Appendix A

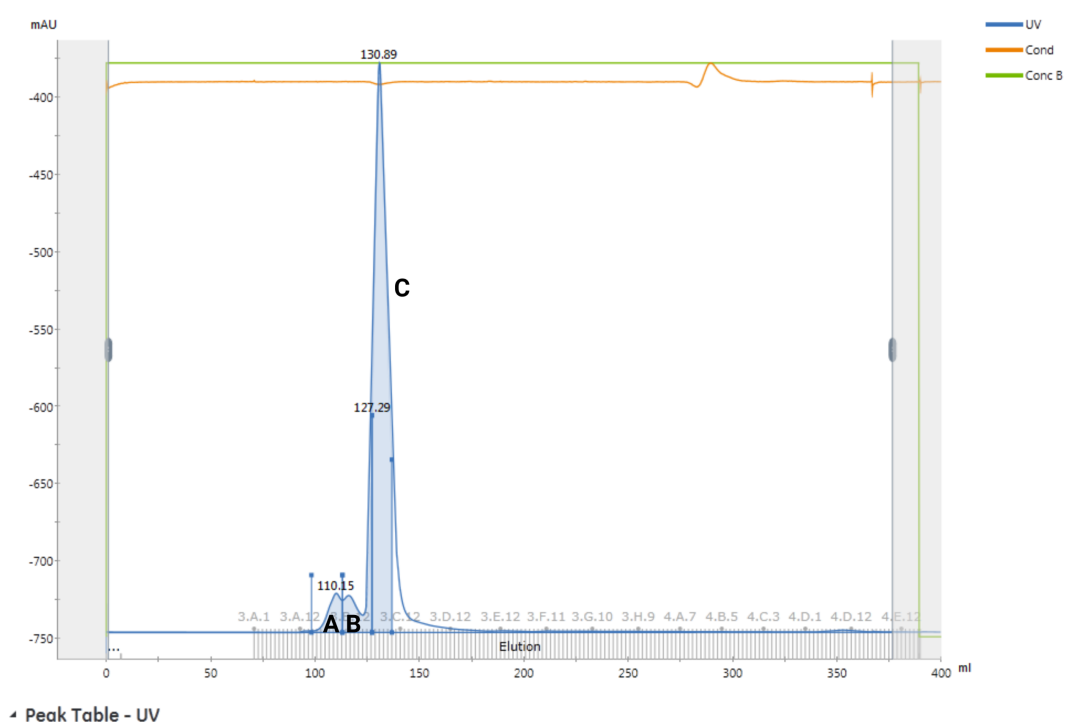


Peak Table - UV

Peak	Retention ml	Area ml*mAU	Area %	Ext coeff. $\text{mg ml}^{-1} \text{cm}^{-1}$	Fraction(s)	Volume ml	Amount mg	Concentration mg/ml	Conductivity mS/cm
Peak A	52.965	170427	91.05		5.A.2 - 5.B.1	104.190			36.06
Peak B	140.378	6030	3.22	0.826	3.B.5 - 3.B.9	9.994	36.502	3.652	35.57
Peak C	190.757	10714	5.72		3.C.1 - 3.E.4	54.546			36.14

**Fig. A.1:** His-Trap HP purification of 6HisAviPaPanK. The first broad peak A represents the loading of the column with lysate. The narrow Peak B represents protein which released from the column after the imidazole concentration increased. This peak was saved for later processing. The flat Peak C occurred at the highest concentrations of imidazole and was not processed further.

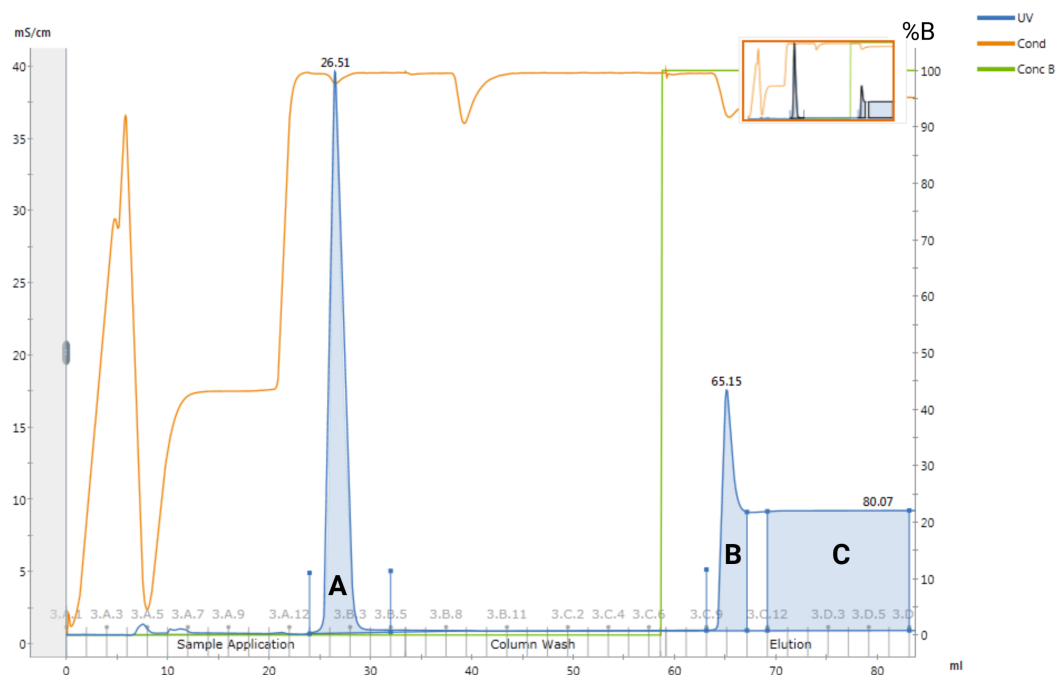




Peak Table - UV

Peak	Retention ml	Area ml* <sup>2</sup> mAU	Area %	Ext coeff. $\text{mg ml}^{-1} \text{cm}^{-1}$	Fraction(s)	Volume ml	Amount mg	Concentration mg/ml	Conductivity mS/cm
Peak A	110.152	183.6	6.04		3.B.2 - 3.B.10	14.813			19.02
Peak B	127.287	378.7	12.45		3.B.10 - 3.C.5	14.187			19.02
Peak C	130.886	2480	81.52	0.826	3.C.5 - 3.C.10	9.545	15.013	1.573	18.96

**Fig. A.2:** Further purification of 6HisAviPaPanK by SEC. Two small peaks A and B eluted just prior to the large peak C. Peak C was expected to contain purified 6HisAviPaPanK and was saved for further analysis.



Peak Table - UV

Peak	Retention ml	Area ml <sup>2</sup> mAU	Area %	Ext coeff. mg ml <sup>-1</sup> cm <sup>-1</sup>	Fraction(s)	Volume ml	Amount mg	Concentration mg/ml	Conductivity mS/cm
Peak A	26.513	1689	26.8	0.677	3.B.1 - 3.B.4	8.001	12.472	1.559	39.40
Peak B	65.148	932.1	14.79		3.C.9 - 3.C.10	3.998			37.77
Peak C	80.066	3680	58.4		3.C.12 - 3.D.6	13.997			37.83

**Fig. A.3:** Reverse His-trap purification following TEV protease cleavage of 6HisTEV *PaPanK*. '%B' on the right y axis denotes the percentage of elution buffer containing 500 mM imidazole is applied to the column. Peak A denotes the sample which eluted during the binding buffer containing 20 mM wash phase and expected to contain tag-cleaved *PaPanK*. Peak B eluted during the isocratic elution as elution buffer containing 500 mM imidazole was applied, and was expected to contain any uncleaved 6HisTEV *PaPanK* and cleaved 6His tags. The flat 'peak' C should contain high amounts of imidazole, which absorbs UV light.

Structure	<i>Pa</i> PanK-B1 (preliminary)	<i>Pa</i> PanK-ATP (preliminary)	<i>Pa</i> PanK-Pant	<i>Pa</i> PanK
<b>Data collection and processing</b>				
Space group	P 41 21 2	P 41 21 2	P 41 21 2	P 41 21 2
a, b, c (Å)	91.85 91.85 180.39	92.45 92.45 179.84	91.24 91.24 180.81	91.98 91.98 181.72
$\alpha, \beta, \gamma$ (°)	90.00 90.00 90.00	90.00 90.00 90.00	90.00 90.00 90.00	90.00 90.00 90.00
Solvent content (%)	66	66	64	65
<b>Diffraction data</b>				
Resolution range (Å)	45.92-1.7	46.22-2.09	81.46-1.75	82.07-2.04
Unique reflections	(1.80-1.70)	(2.22-2.09)	(1.89-1.75)	(2.07-2.04)
Multiplicity	161265 (25532)	46540 (7243)	63310 (3166)	49562 (2476)
R meas (%)	7.85 (7.79)	14.71 (14.92)	13.1 (13.4)	12.8 (13.8)
Completeness (%)	15.3	15.1	11.1	11.0
I/sigI	99.6 (97.5)	99.6 (97.7)	94.3 (57.1)	98.6 (100)
	7.83 (-0.02)	14.11 (1.21)	14.4 (1.7)	14.8 (2.1)
<b>Refinement</b>				
R work/R free	0.19/ 0.22	0.22/0.25	0.17/0.20	0.15/0.20
Quaternary structure	dimer	dimer	dimer	dimer
Protein residues (in a dimer)	486	485	495	486
Cations	-	1 * Mg2+	2 * Ca2+	-
Ligands	1 * B1	1 * ATP	2 * Pantothenic acid	-
Water molecules	246	-	257	211
<b>Rmsds</b>				
Bonds (Å)	0.016	0.009	0.002	0.004
Angles (Å)	1.97	1.74	1.19	1.24
<b>Ramachandran plot, residues in (%)</b>				
Favored regions	99.38	96.91	99.39	99.59
Allowed regions	0.21	2.47	0.20	0
Outlier regions	0.41	0.62	0.40	0.41
<b>Mean B factors (Å<sup>2</sup>)</b>				
Protein atoms	58.73	58.63	39	57.94
Ligands	81.04	78.58	33	-

**Table A.1:** Collected data and refinement statistics following x-ray diffraction of *Pa*PanK crystals soaked with various ligands. Data in parentheses pertain to the highest resolution shell.