

Understanding the development of industrial strains of *Streptomyces*

By

Molly Keith-Baker

Primary supervisor: Professor Paul. A Hoskisson

Secondary supervisor: Professor Iain. S Hunter

**A thesis submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy at the University of Strathclyde.**

Strathclyde institute of Pharmacy and Biomedical Sciences

October 2023

Declaration

'This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.'

'The copyright of this thesis belongs to the author under the terms of the United Kingdom Copyright Acts as qualified by the University of Strathclyde Regulation 3.50. Due acknowledgement must always be made of the use of any material contained in, or derived from, this thesis.'

Signed:  Date: June 2025

COVID impact statement.

I started my PhD in October 2019 – six months before the COVID-19 pandemic resulted in a UK-wide lockdown and the complete and immediate closure of the University. As a result, I had no access to the lab during the early stages of my PhD project. While the labs were closed for a period of 5 months (reopening August 2020), the expected lengths of the closures were unclear at the time. This was a very challenging experience as there was limited work that I could complete from home and the uncertain nature of the lockdown made it difficult to plan lab work for the eventual reopening of the University. Once access to the lab was granted in August 2020, it was far from normal operating conditions, with working hours restricted to 50% of the pre-pandemic conditions and limited out of hours working to minimise mixing of people and risk of transmission. While this was a necessary decision in terms of pandemic control, it made certain experiments impractical and logistically unfeasible, and ultimately slowed progress in my project. In addition, BBSRC decided against granting extensions to my cohort (due to us being so early in our PhDs) which meant some of delayed and disrupted experiments had to be abandoned entirely in order to deliver the project on the original timeline.

Acknowledgements

I would like to thank my examiners Dr Arnaud Javelle and Professor Paco Barona Gomez for taking the time to read and examine my thesis. Thank you also to IBioiC and the BBSRC for funding my PhD and to our collaborators at GSK, Andrew, Ben, Nicky, Steve, and Ted for their interest, ideas, and support during our meetings. I would also like to thank Professor Gilles Van Wezel for giving us the CRISPR plasmids used throughout this thesis.

First and foremost, I would like to thank my primary and secondary supervisors Professor Paul Hoskisson and Professor Iain Hunter. How lucky I have been to have such great scientists supporting me. Paul thank you for giving me the chance to do a PhD, Dave was right when he told me you'd be a great supervisor. You have supported me through a pandemic as well as a PhD, and being in such an encouraging environment has allowed me to thrive and hopefully become a scientist you can be proud of. Iain thank you for all your words of wisdom and encouragement I always feel like I come away from a conversation with you having learnt something new.

I would also like to thank Professor David Whitworth my undergraduate and masters supervisor for being the first academic to really make me feel like they believed in me, and without whom I wouldn't have made it to Glasgow.

A huge thank you to everyone from HW601 for bringing the laughs and support not just the science, I've have been so truly lucky that my colleagues have become such great friends over the past four years.

A special thanks to all the Hoskie group members past and present, it has been so lovely to be a part of this group, but an extra mention has to go to our wonderful post docs who really have made finishing this PhD that little bit easier.

To 'The gals' I knew from the moment I met you at GMC in 2019 that we'd be the best of friends. Thank you to all the level 6 gals for all the coffees, pints, laughs, dinners, nights out, nights in, and trips away over the last four years, you really have made my PhD experience.

Aberystwyth was not only where I started my academic journey, but where I had four wonderful years and met three of the best friends a girl could ask for. To Charlie, Josh, and Meg thank you for all the encouragement, support, and weekends away that have allowed me to escape over the last four years, you mean the world to me.

Gordon, the past four years have not only brought me (hopefully) a PhD, countless wonderful friends, and some of the best memories, but they also brought me you. You have treated me with nothing but love, support, and compassion, more than I have sometimes deserved, and for that I will never be able to thank you enough. You inspire me every day to be a better person and a better scientist, and have taught me so much over. I love you and will always endeavour to show you the same love and support that you continue to show me. Every oxford comma in this thesis is dedicated to you.

To my Grandparents, who I love dearly. Thank you for always showing an interest even if you didn't understand what I was talking about. I'm sorry that you're not all here to see this.

I have saved the best until last, my Mum and Dad. Thank you for always supporting my dreams, being my biggest cheerleaders, and the ones I know I can always rely on. I hope I've made you proud. My a'th kar.

‘Humans only have one ending. Ideas live forever.’

- Ruth Handler

Table of Contents

<i>Declaration</i>	2
<i>COVID impact statement</i>	3
<i>Acknowledgements</i>	4
<i>Abstract</i>	12
<i>List of abbreviations:</i>	14
<i>Chapter 1: Introduction</i>	15
1.1 Actinobacteria and Natural Products	16
1.2 Antibiotics and the risk of AMR	17
1.3 Streptomyces	20
1.3.1 <i>Streptomyces coelicolor</i>	23
1.3.2 <i>Streptomyces clavuligerus</i>	24
1.4 Regulation	24
1.4.1 Regulation of antibiotic biosynthesis by metabolism: Carbon catabolite repression	25
1.4.2 Regulation of Actinorhodin biosynthesis.	26
1.5 Biosynthesis of Actinorhodin.	27
1.6 Regulation of Clavulanic acid biosynthesis.	27
1.7 Biosynthesis of Clavulanic acid.	30
1.8 Clavulanic acid mode of action	33
1.9 The GlaxoSmithKline <i>Streptomyces clavuligerus</i> lineage.	35
1.10 Microbial Metabolism	37
1.11 Central node of metabolism	37
1.12 Pyruvate phosphate dikinase	42
1.14 PPKD multiplicity	44
1.15 Aims:	48
<i>Chapter 2: Materials and methods</i>	49
2.1 Bioinformatics	50

2.1.1 Multiple sequence alignments	50
2.1.2 Phylogenetic trees.....	50
2.1.3 Predictions of protein structures	50
2.2 Microbiology	51
2.2.1 Bacterial strains.....	51
2.2.2 Media	54
2.2.3 Buffers	57
2.2.4 Antibiotics	60
2.2.5 Culture conditions	61
2.2.6 Spore preparation.....	61
2.2.7 Spore quantification.....	61
2.2.8 Growth curves using Cell Growth Quantifier (CGQ) [Aquila biolabs]	62
2.2.9 Interspecies complementation experiments.....	62
2.2.10 Cell dry weight	63
2.2.11 Actinorhodin assay	63
2.2.12 Screening <i>Streptomyces</i> for phenotypes	64
2.2.13 Biolog	64
2.3 Molecular biology	66
2.3.1 Plasmids.....	66
2.3.2 Preparation of chemically competent <i>E. coli</i>	68
2.3.3 Plasmid miniprep.....	68
2.3.4 PCR clean up and Gel extraction	68
2.3.5 Genomic DNA extraction	69
2.3.6 RNA extraction.....	69
2.3.7 PCR reaction for sgRNA amplification	70
2.3.8 Colony PCR for <i>E. coli</i>	70
2.3.9 RT-PCR	70
2.3.10 Oligonucleotide primers	72
2.3.11 Agarose gel electrophoresis	73
2.3.12 Restriction digest	73
2.3.13 Restriction cloning	73
2.3.14 TA cloning	73

2.3.15 Transformation of <i>E. coli</i> by heat shock	74
2.3.16 Intergenic conjugation	74
2.4 Protein expression	75
2.4.1 Cell lysis and preparation for purification	75
2.4.2 Immobilised Metal Affinity Chromatography (IMAC)	76
2.4.3 SDS-Poly Acrylamide Gel Electrophoresis	76
2.4.4 PPDK enzyme activity assay	77
Chapter 3: Generation of <i>ppdk</i> mutants in <i>S. coelicolor</i> and <i>S. clavuligerus</i> using CRISPRi technology.....	78
3.1 Introduction:	79
3.2 Aims and objectives of this Chapter:	81
3.2.1 Specific objectives:	81
3.3 Designing guide RNA and generating CRISPR vectors.....	83
3.4 PCR to introduce 20 bp protospacer into sgRNA construct	83
3.5 Cloning of sgRNA construct to generate final CRISPR vectors	89
3.6 Synthesis of <i>ppdk</i> over expression plasmids	91
3.7 Conjugation of knockdown and overexpression plasmids into <i>Streptomyces spp.</i>	93
3.8 RT-PCR used to confirm knockdown of <i>ppdk</i> genes	95
3.9 Discussion:	97
Chapter 4: Changing the expression of <i>ppdk</i> alters the physiology and growth characteristics of <i>S. coelicolor</i> and <i>S. clavuligerus</i>	99
4.1 Introduction	100
4.2 Aims and objectives of this Chapter:	101
4.2.1 Specific objectives:.....	102
4.3 Analysis of nutrient utilisation using phenotypic Biolog arrays shows decreasing substrate utilisation in industrial <i>Streptomyces spp.</i>	103
4.4 Investigating growth rates using single carbon sources.....	109

4.5 Altering expression of <i>ppdk2</i> (SCO2494) in <i>S. coelicolor</i> impacts antibiotic titre	118
4.7 Discussion	127
4.7.2 Growth analysis of strains.....	129
4.7.3 Physiological role of <i>ppdk</i>	130
Chapter 5: Expression, purification, and characterisation of Pyruvate phosphate dikinase proteins from <i>S. clavuligerus</i> and <i>S. coelicolor</i>..	132
5.1 Introduction	133
5.2 Specific Objectives:	134
5.3 PPDK sequence is conserved throughout the industrial lineage of <i>S. clavuligerus</i>	135
5.4 Comparison of <i>Streptomyces</i> PPDK sequences and sequences of previously characterised PPDK proteins.....	137
5.5 Alignment shows PPDK structures are conserved throughout different kingdoms of life.....	140
5.6 Structures of PPDK proteins from <i>S. clavuligerus</i> and <i>S. coelicolor</i> were predicted using AlphaFold2.	142
5.7 Construction of vectors for the overexpression of Streptomycete PPDK proteins.	148
5.8 Overexpression of Streptomycete PPDK proteins.....	151
5.9 Optimisation of PPDK purification	153
5.9.1 Gradient elution unexpectedly reduced yield of protein	156
5.9.2 Purifying PPDK proteins through step elution increases yield of protein.....	163
5.10 Enzymatic characterisation of purified PPDK proteins.	170
5.11 Discussion	179
Chapter 6: Discussion	182
6.1 General discussion	183
6.1.1 Traditional strain development reduces carbon utilisation profile in bacterial strains.....	184
6.1.2 PPDK performs a unique role in mediating carbon flux in <i>Streptomyces spp.</i>	186

6.1.3 PPK proteins in <i>Streptomyces spp</i> are canonical with PPKs characterised in other genera	188
6.2 Conclusion	190
References:	194
Appendix 1:	216

Abstract

Due to the current AMR crisis there is a need to produce more antimicrobials, however discovering new antibiotics is a long and expensive process. As a result, there is also a need to increase the production of current antimicrobials and specialised metabolites.

Previous work has shown that in *Streptomyces coelicolor* the primary metabolic enzyme Pyruvate phosphate dikinase (PPDK), which catalyses the reversible reaction of pyruvate to phosphoenolpyruvate (PEP), is up-regulated almost 30 fold during antibiotic production on oil based media. This thesis further investigated this observation by studying pyruvate PPDK and its role in facilitating specialised metabolite production. PPDK like Pyruvate kinase (Pyk) can convert PEP to Pyruvate, however the reaction is reversible and the PEP forming reaction is favoured. Actinobacteria are relatively unique having this enzyme, as it is uncommon among bacteria and it importantly enables the organism to grow gluconeogenically as well as glycolytically. Several genes encoding primary metabolic enzymes show multiplicity in *Streptomyces* and there are commonly two PPDK proteins. However, the two versions of this enzyme are in separate clades when mapped onto a phylogenetic tree, supporting the hypothesis that the two may have different physiological roles.

Using CRISPRi technology a number of *ppdk* knockdown and over expression mutants were produced and their carbon utilisation profiles were analysed. These preliminary data were the basis for analysing growth rate and antibiotic titre on single carbon sources, which showed that by manipulating these central carbon metabolic genes in *Streptomyces* it is possible to alter antibiotic titre. Further bioinformatic and biochemical analysis of these enzymes also shed light on their physiological roles in *Streptomyces*. These results also help to further understand carbon flux around the PEP-PYR-OXA (PEP-pyruvate-oxalacetate) node of central carbon metabolism under

a range of conditions and to understand how this impacts the availability of precursor molecules for specialised metabolite production.

List of abbreviations:

ACT – Actinorhodin

AMR – Antimicrobial resistance

BGC – Biosynthetic gene cluster

CA – Clavulanic Acid

CCR – Carbon catabolite repression

CDA – Calcium dependent antibiotics

CDS – Coding sequence

CSR – Cluster situated regulator

G3P – Glyceraldehyde-3-phosphate

HGT – Horizontal gene transfer

MBL - Metallo- β -lactamases

MFS – Major facilitator superfamily

MSA – Multiple sequence alignment

NJ – Neighbour joining

OM – Outer membrane

ORF – Open reading frame

PCD – Programmed cell death

PPP – Pentose phosphate pathway

PTS – Phosphotransferase system

SARP – Streptomyces antibiotic regulatory protein

TCA – Tricarboxylic acid

Chapter 1: Introduction

1.1 Actinobacteria and Natural Products

Natural products are important because they encompass products such as anticancer, antiviral, and antibiotic drugs. Bacteria and other microorganisms such as fungi are a great source of these specialised metabolites and natural products which are used regularly in clinical settings and are of great commercial value to medical and agricultural industries [1].

The *Actinomycetota* are one of the largest phyla within the bacterial domain, they are mostly free-living organisms and inhabit marine, freshwater, and terrestrial environments [2]. Not only are Actinobacteria ubiquitous but they also inhabit some of the most extreme of the above environments including arid environments such as the Atacama Desert [3]. They are also present in sediment of the Mariana Trench, which exposes these organisms to extremes of, light, salinity, and pressure [4]. Cultivating Actinobacteria from these extreme environments has shown great potential with regards to the discovery of natural products. The Actinobacteria sequenced thus far have proved to be invaluable to the areas of both human and veterinary medicine as well as informative in terms of their biotechnological potential [2]. Although Actinobacteria have a wide range of habitats, soil environments are home to a large proportion of the species. They produce specialised metabolites, for example enzymes that can be utilised to access nutrients and siderophores for interspecies interaction [5, 6]. It is not just Actinobacteria that utilise their specialised metabolites in this way, *Myxobacteria* swarm their 'prey' and expel lytic enzymes to breakdown other organisms releasing nutrients into the surrounding environment [7].

For several decades Actinobacteria have been amongst the most important and extensive sources of novel antibiotics, many of which are still used clinically. These include a number of different classes of drugs including β -lactams, tetracyclines, rifamycins, aminoglycosides, macrolides and glycopeptides [8]. While Actinobacteria

produce more than half of all clinically relevant antibiotics and many other life-saving secondary metabolites, the phylum also contains many deadly pathogens. This includes *Mycobacterium* and *Corynebacterium* species which can cause tuberculosis and diphtheria in humans, making them species of interest in the field of medical research [2] [9].

1.2 Antibiotics and the risk of AMR

The World Health Organisation has recognised antimicrobial resistance (AMR) as one of the biggest threats to global health, which is why it is imperative to continue research into new antibiotics and increasing yield of antimicrobials currently in the clinic.

There are four main modes of action of antibiotics: inhibition of metabolic pathways, inhibition of proteins synthesis, interference in nucleic acid synthesis, and interference with cell wall synthesis. The lattermost is potentially the most well known as this is the underlying mechanism by which β -lactam antibiotics such as penicillins work [10, 11]. β -lactams are the most commonly used class of antibiotics and include penicillins, cephalosporins, and carbapenems. The mechanism by which β -lactams kill pathogens is by targeting penicillin binding proteins, which are enzymes that inhibit the transacylation reaction. This is the last step in the biosynthesis of the peptidoglycan cell wall, and results in death of the growing bacteria.

Irrespective of antibiotic class or mechanism of action, there is a risk of antimicrobial resistance emerging. Bacteria can be intrinsically resistant to antibiotics, they can acquire resistance via genetic mutations that increase their fitness or they can acquire resistance mechanisms via horizontal gene transfer [12, 13]. One example is that Gram-negative bacteria are intrinsically resistant to many antibiotics that commonly

inhibit Gram-positive bacteria because of their outer membrane (OM). This is composed of lipids and proteins which acts as a tight molecular sieve meaning that only very small molecules can penetrate the cell. However, this does mean that these Gram-negative bacteria require other mechanisms by which to transport nutrients into the cell, often this is by incorporating porin channels into the OM [9–11].

The acquired mechanisms of resistance can be split into three categories: target modification, drug transport, and drug inactivation [14]. Organisms can possess resistance by post translational modifications, an example of this is methylation of ribosomal proteins which can prevent the binding of aminoglycosides which inhibit protein biosynthesis. Drug transport is another mechanism of resistance by which pathogens avoid drugs coming into contact with their target. A way of doing this is using efflux pumps to actively transport antimicrobials that have entered the cell back out, an example is NorA a member of the major facilitator superfamily (MFS) used by *Staphylococcus aureus* [15, 16]. The third mechanism of resistance is by inactivation of the antibiotic either via hydrolysis or steric hindrance. This mechanism is of particular importance because it is one of the most prevalent resistance mechanisms used by pathogens to evade β -lactam antibiotics. The β -lactamase enzymes hydrolyse the amide bond in the β -lactam ring resulting in deactivation of the molecule and this is a mechanism used by both Gram-positive and Gram-negative pathogens alike [17, 18]. There are four classes of β -lactamase enzyme (A-D), three of which (A, C, and D) are active site serine enzymes. These catalyse a three-step reaction, first forming a non-covalent complex with the β lactam, then acetylating the active site serine side chain, and finally the complex undergoes deacetylation and hydrolysis, the enzyme is then regenerated. These all show sequence similarity and have very similar structures suggesting that these active site serine enzymes have the same evolutionary origin [12, 19]. Class B β -lactamases also known as metallo- β -lactamases (MBLs) require a bivalent metal ion, often zinc, for their activity and are

structurally different from the active serine enzymes. MBLs show resistance against most β -lactamase antibiotics and utilise the active site Zn^{2+} ions to activate an active site water molecule used to hydrolyse the β -lactam ring [20, 20, 21].

1.3 *Streptomyces*

The genus *Streptomyces* is amongst the most speciate bacterial genera with >700 validly described species. It is of particular interest as it encompasses bacteria that are functionally diverse producing an excess of enzymes that enable them to survive in a wide range of environments where they play a key role in carbon cycling [22, 23]. This role is facilitated by their production of a multitude of secreted hydrolytic enzymes and the production of an abundance of specialised metabolites [2]. They have a complex lifecycle illustrated in Figure 1.1 and exhibit many phenotypic traits that are uncommon in other bacteria [24]. The *Streptomyces* lifecycle begins with a single spore that germinates to form a mat of vegetative hyphae which forms the main body of the colony. Under stressful conditions, for example a diminished source of nutrients, the mycelia begin to differentiate into substrate and aerial hyphae, the aerial hyphae grow up into the air and undergo synchronous septation forming chains of unigenomic spores. This differentiation stage of the lifecycle coincides with the production of antibiotics and other secondary metabolites [2, 24–26].

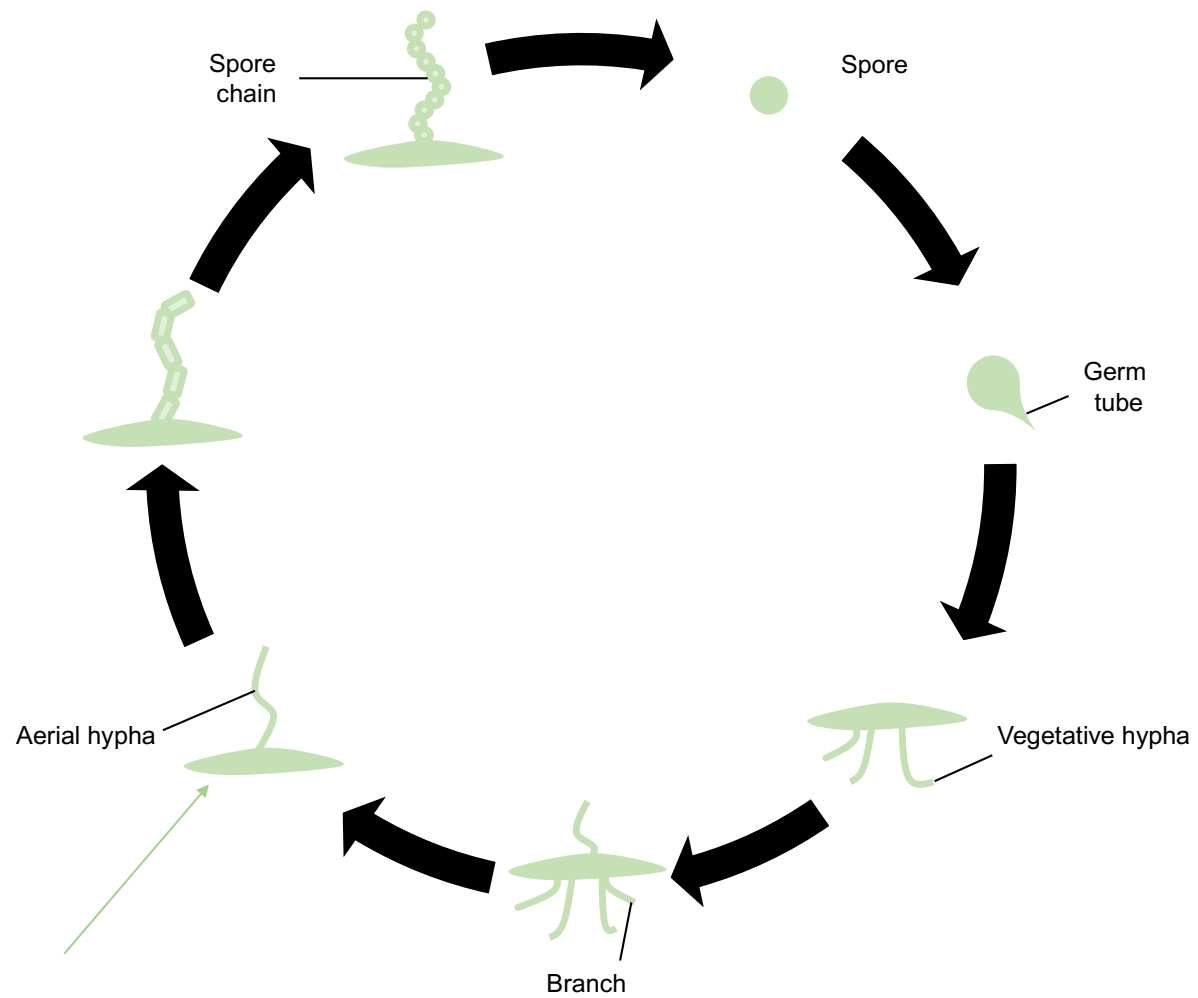


Figure 1.1 Lifecycle of *Streptomyces*, green arrow indicates the stage at which antibiotics are produced.

Sporulation is intimately linked to the production of specialised metabolites, however, a number of species are unable to sporulate when grown in liquid culture in a laboratory setting. This makes it difficult to study the secondary metabolite production as this stage of the lifecycle halts metabolism [25]. It has also been shown that metabolite production is triggered when nutrient availability is limited with examples of carbon, nitrogen and phosphate inducing bioactive metabolite production [26, 27]. When nutrients become depleted, the substrate and vegetative mycelia of Streptomyces are hydrolysed and degraded by a mechanism similar to programmed cell death (PCD) to release nutrients (such as amino acids, sugars, and lipids) required for the colony to complete the developmental process [2]. In most species of Streptomyces, differentiation and secondary metabolism are co-regulated and it is thought that antibiotics might be produced at this stage to protect the newly released nutrient pool from competitor microorganisms. However there has also been evidence that secondary metabolites produced at this stage are involved in communication and signalling [2, 28, 29]. The roles played by natural products of Streptomyces enable them to shape their environment and communities, for example symbiosis and environmental interactions [30], interactions with ants [31–33], wasps [34], marine organisms [35, 36], and plants [37–39]. Production of specialised metabolites can also enable collective traits (nutrient scavenging, antibiosis and signaling) but this occurs at the expense of using growth precursors, cofactors and energy [40].

The exploitation of Streptomyces for secondary metabolite production began in 1942 with the discovery of streptothricin, followed by streptomycin in 1944. Since then, research into Streptomyces for antibiotics intensified and the genus still provides over 80% of antibiotics [28]. Although they have been used for their secondary metabolites for almost a century, Streptomyces are still a rich source of novel clinically relevant metabolites. Such compounds are invariably the products of the large number of complex biosynthetic gene clusters (BGCs) present in Streptomyces

genomes [41]. BGCs are co-located groups of genes, which encode enzymes, regulators, and transporters for a given metabolic pathway.[42].

Prior to the advent of cheap genome sequencing technologies, most species of *Streptomyces* were believed to produce only one or two bioactive molecules so were expected to possess a small number of BGCs. However recent genome analysis revealed an excess of previously unknown clusters, termed 'silent BGCs' [41]. There is currently a large effort to express these clusters in heterologous *Streptomyces* host strains, however this is a difficult process and trying to express silent BGCs under laboratory conditions has proven challenging. This abundance of unelucidated silent BGCs suggests a wealth of specialised metabolites that are yet to be discovered [43].

1.3.1 *Streptomyces coelicolor*

There are several *Streptomyces* species routinely used as a heterologous host. These include: *Streptomyces coelicolor*, *Streptomyces venezuelae*, and *Streptomyces lividans*. *S. coelicolor* has been considered a model organism for over 50 years. It is well characterised and easy to genetically manipulate compared with some other Streptomycetes [44]. *S. coelicolor*, like most other Streptomycetes, is characterised by a complex lifecycle, as described previously, and its ability to produce a variety of secondary metabolites. This model organism is known to produce a number of natural products including actinorhodin, undecylprodigiosin, and calcium dependant antibiotics (CDA), the production of which are linked to the morphological differentiation that happens during the lifecycle [45, 46].

1.3.2 *Streptomyces clavuligerus*

Streptomyces clavuligerus was discovered in the early 1970s [47]. It produces a number of specialised metabolites including cephamycin C and clavams, most notably clavulanic acid, which are β -lactam ring containing specialised metabolites [48]. Clavulanic acid (CA) is of particular importance because it has a strong inhibitory effect against β -lactamases which are produced by pathogens as a defence to degrade β -lactam antibiotics [49]. The biosynthetic pathways that produce clavulanic acid and cephamycin C are located in a supercluster, although both molecules are synthesised from distinct precursor molecules [50].

1.4 Regulation

A number of regulatory mechanisms, acting at all levels of expression, have been identified for specialised metabolite production in *Streptomyces*. These range from transcriptional control at the single gene level through to translational control mechanisms [51]. Direct control of BGCs was first identified through pathway specific regulator proteins (now referred to as Cluster Situated Regulators [CSR]) [52] such as ActII-ORF4, RedD, CdaR and StrR [53–56]. These were amongst the first members of the well-studied *Streptomyces* antibiotic regulatory protein (SARP) family to be identified [56]. These were later renamed CSRs, following the discovery of regulators that could act on biosynthesis beyond a single, specific pathway. Later the roles of extracellular signalling molecules, such as the γ -butyrolactones, were identified as inducers of specialised metabolites through the binding of cytoplasmic binding proteins, which activate CSRs in a range of species [57–59]. The identification of global acting pleiotropic regulators of specialised metabolites such as the AfsR and

AbsA1A2, which belong to the two-component regulatory systems [51, 60] in *S. coelicolor*, indicated that there is a complex multi-level control of biosynthesis and highlights that these organisms must integrate a number of signals to regulate specialised metabolite biosynthesis in an appropriate manner [61–65]. This work has been reviewed extensively [29, 66] however recent work has added greater detail and new mechanisms of regulation have emerged, reviewed by Hoskisson & Fernandez-Martinez in 2018 [27].

1.4.1 Regulation of antibiotic biosynthesis by metabolism: Carbon catabolite repression

In *Streptomyces*, the main mechanism used to control different physiological processes is carbon catabolite repression (CCR). This mechanism controls the genes involved in the uptake and utilisation of different carbon sources, depending on whether more favourable carbon sources are present in the culture media. This mechanism is largely independent of the phosphoenolpyruvate phosphotransferase system (PTS) [67]. PTS is a common mechanism amongst bacteria used for the uptake of different carbohydrates across the bacterial cell envelope [68]. In most bacteria, CCR is mediated by the phosphoenolpyruvate-dependent phosphotransferase system (PTS), yet in *Streptomyces*, glucose uptake is mediated by the Major Facilitator Superfamily (MFS) transporter, glucose permease (GlcP), and there is evidence for direct interaction between glucose kinase (Glc) and GlcP which may mediate CCR [69]. The role of nutrient-sensing in regulation of antibiotic biosynthesis is well known [70] with the enzyme glucokinase (Glc) playing a central role in carbon-catabolite repression in *Streptomyces* [71].

Antibiotic synthesis is highly regulated by not only specific genes but also the availability of certain nutrients, like carbon, nitrogen, and phosphates. CCR is initiated

when there are high concentrations of carbon sources available, e.g. glucose can cause the down regulation of secondary metabolism, slowing or even halting completely the production of antibiotics and other specialised metabolites. Many reports highlight glucose as a major mediator of this mechanism, suppressing secondary metabolism by favouring primary metabolism and thus reducing the available pool of precursors such as glyceraldehyde-3-phosphate (G3P). Although glucose is often the initiator of this mechanism other transcriptional regulators in *Streptomyces* have emerged as having potential roles in the mechanism. Glycerol also induces CCR in a number of species: *S. parvulus*, *S. garyphalus*, and *S. clavuligerus*, where it causes the repression of actinomycin, cycloserine, and cephamycin c synthase production respectively [67].

This suggests that there is not yet a clear understanding of how glucose-mediated CCR in *Streptomyces* functions in relation to specialised metabolite precursor supply.

1.4.2 Regulation of Actinorhodin biosynthesis.

Actinorhodin biosynthesis in *S. coelicolor* is controlled by a Cluster Situated Regulator (CSR), ActII-ORF4. This transcription factor binds directly to genes within the BGC and plays a pivotal role in activating transcription [29, 72]. ActII-ORF4 is negatively regulated by the AbsA1- AbsA2 two-component system. AbsA1 is the sensor kinase which phosphorylates AbsA2 and subsequently binds to the DNA target to down regulate Actinorhodin production. This two-component system also regulates production of undecylprodigiosin, and other calcium dependent antibiotics produced by *S. coelicolor* [29].

1.5 Biosynthesis of Actinorhodin.

In *S. coelicolor* seven molecules of acetyl-CoA and one molecule of malonyl-CoA are required for the biosynthesis of actinorhodin (ACT). The rate of carbon flux through glycolysis and towards the tricarboxylic acid (TCA) cycle impacts how much acetyl-CoA is available as a precursor for actinorhodin and for undecylprodigiosin production. It has been shown that channelling carbon in this direction rather than allowing it to enter the pentose phosphate pathway (PPP), by deleting *zwf1/2*, which encodes glucose-6-phosphate dehydrogenase, increases the level of ACT production in *S. lividans* [73].

1.6 Regulation of Clavulanic acid biosynthesis.

Although Cephamicin and CA are both β -lactams produced by *S. clavuligerus* they are, as mentioned previously, derived from different precursors 2-oxoglutarate and lysine, and G3P and arginine respectively [74]. The production of cephamicin C is repressed by the presence of glycerol, which simultaneously increases the biosynthesis of CA [75]. Wild-type *S. clavuligerus* is unable to use glucose as a carbon source, however glycerol is readily taken up and is a good carbon source for CA production. Growth on glycerol ensures a sufficient supply of glyceraldehyde-3-phosphate (G3P) for CA biosynthesis. Glycerol is metabolised by glycerol dehydrogenase, aldehyde dehydrogenase and glycerate kinase to produce 2-phosphoglycerate which enters glycolysis. A knockout of glycerate kinase is proposed to have potential to redirect the anaplerotic pathways towards G3P, decreasing the amount of C3 precursor directed towards amino acid biosynthesis and therefore enabling the increase of C3 pool for CA production [76]. Due to the fact *S. clavuligerus*

is able to utilise glycerol efficiently, oils are extensively used in industry as a source of glycerol from mono and triacylglycerols for CA fermentation [77]

The regulation of CA biogenesis can be broken down into two parts: global regulation and pathway specific regulation [78]. Global regulators are not located within the BGC and therefore often have pleiotropic effects, not only controlling antibiotic production but also morphology of the organism, the flux within primary metabolism and response to stress signals - for example nutrient deficiency [79, 80]. γ -butyrolactone is one global regulator in *Streptomyces* that can trigger pathway specific regulation. CSRs are lower level regulators that reside inside the BGC and directly regulate transcription [79]. There are several types of CSR but the two involved in the regulation of CA biosynthesis are (1) CcaR, a *Streptomyces* Antibiotic Regulatory Proteins (SARP) which is a regulator for not only CA production but also Cephameycin C biosynthesis. It is also known that CcaR directly regulates transcription of the 'early' CA genes (Figure 1.2) and indirectly regulates the transcription of the 'late' genes by positively controlling the other CSR (2), ClaR, a LysR family regulator that positively regulates CA biosynthesis [66, 78–80].

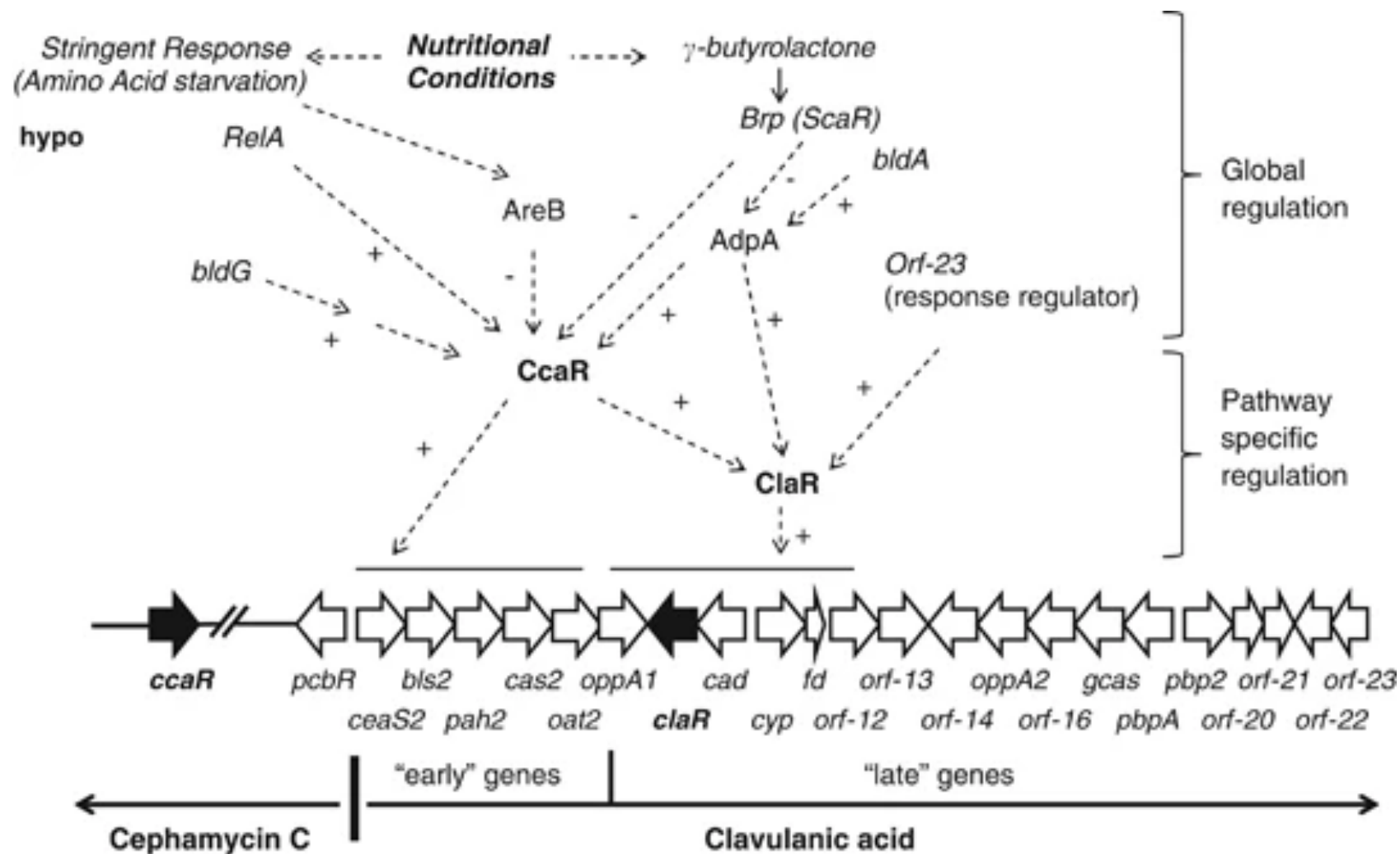


Figure 1.2 Biosynthetic gene cluster of clavulanic acid and its regulators [78].

1.7 Biosynthesis of Clavulanic acid.

CA is produced during a condensation reaction between the glycolytic intermediate G3P and an arginine molecule [78]. As well as CA, *S. clavuligerus* produces a number of other clavam molecules including both cephamycin and penicillin. These molecules are 5S clavams because of their 5R stereochemistry whereas CA has a 3R configuration [81]. Unlike the 5S clavams, CA does not require a tri-peptide precursor molecule; the nitrogen is provided to the β -lactam ring by either arginine or ornithine [82]. For many years the synthesis of the β -lactam ring in CA and cephamycin C was thought to be shared. However it was subsequently demonstrated that cephamycin C uses 2-oxoglutarate and lysine as precursor molecules [74]. CA shares the beginning of its biosynthetic pathway with the 5S clavams, up until the production of clavaminic acid where the pathways diverge, although the biosynthetic machinery are located in distinct BGCs (Figure 1.3).

While cephamycin production is highly reliant on the availability of nitrogen, CA production instead relies heavily on the availability of the precursor molecule G3P [83]. Specialised metabolites are usually produced in nutrient limited environments including the limited availability of phosphate. This may reduce the availability of GAP which is an essential precursor for secondary metabolism and reportedly a limiting factor in CA biosynthesis [49]. Glycerol is metabolised to G3P which can then enter a number of pathways, the TCA cycle, it can be recruited into the gluconeogenesis pathway or finally can be combined with arginine in the CA biosynthetic pathway [84]. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) which is encoded by the *gap* gene, is a key rate limiting step in glycolysis. Experiments in which *gap* is knocked out show an increase in the production of CA, suggesting that there is a larger pool of G3P available as a precursor for CA production. The titre of CA was increased again when the *gap-1* mutant was supplemented with arginine, suggesting that in this case

the pool of G3P had increased to such an extent that arginine had replaced G3P as the limiting factor [84].

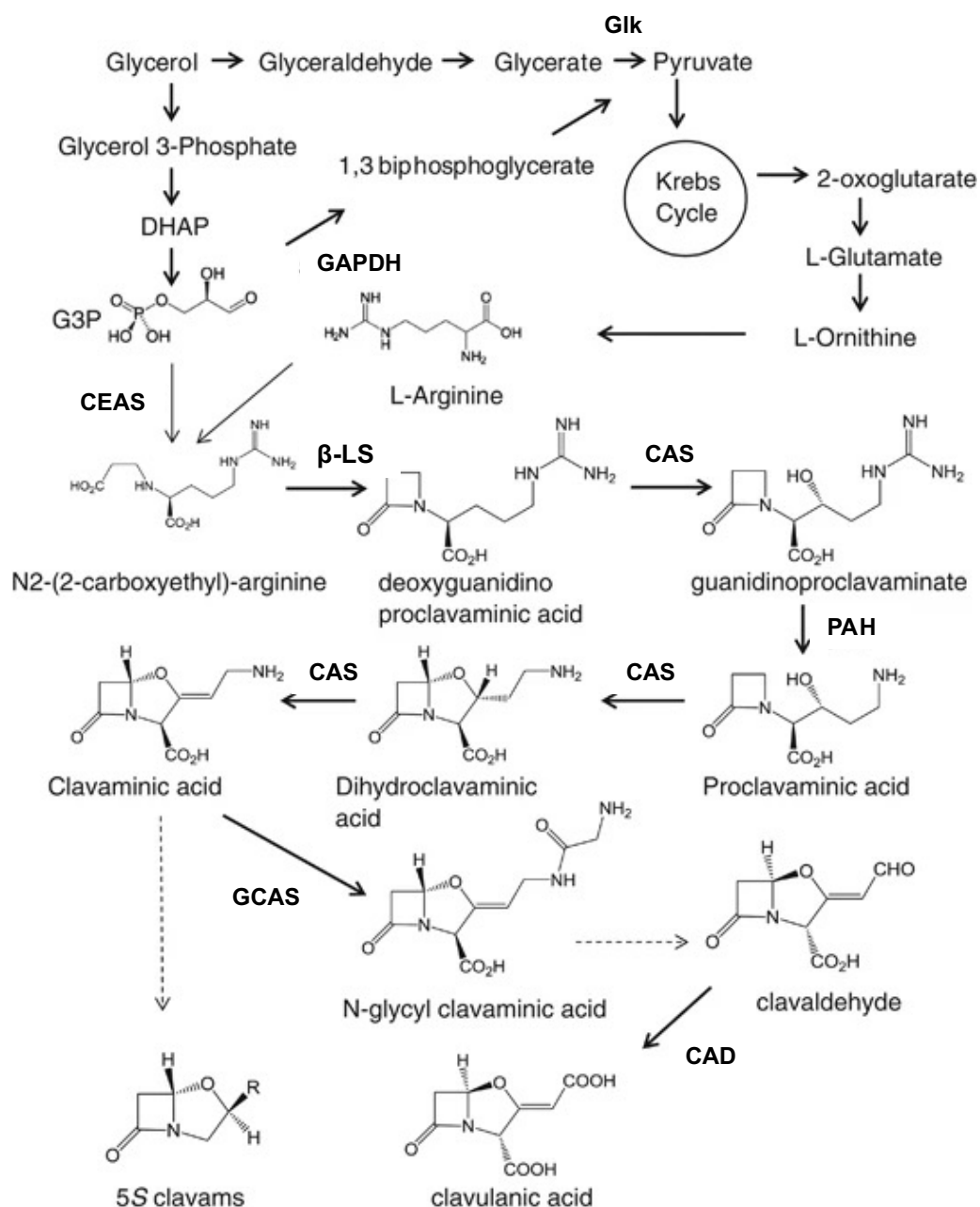


Figure 1.3 Biosynthetic pathway for clavulanic acid production. Showing the precursors glyceraldehyde-3-phosphate and arginine from primary metabolism [78]. Enzymes involved in this pathway include; glycerol kinase (Glk), glyceraldehyds-3-phosphate dehydrogenase (GAPDH), N(2)-(2-carboxyethyl)arginine synthase (CEAS), β-Lactam synthase (β-LS), clavamate synthase (CAS), proclavamate amidinohyrolase (PAH), clavulanic acid dehydrogenase (CAD), and *N*-glycyl-clavaminic acid synthase (GCAS).

1.8 Clavulanic acid mode of action

Although *Streptomyces* are an excellent source of novel antimicrobials such as β -lactams, as previously mentioned the utility of these molecules is under constant threat due to the emergence of resistance in the clinic.

Antibiotic potentiators in this case are particularly important as they are prescribed with existing antibiotics to enhance the activity of the antibiotic. There are a number of ways by which an adjuvant can increase the activity of an antimicrobial, they can enhance uptake of the drug, block efflux pumps exporting the drugs, change the physiology of the target cells, or inhibit the resistance mechanism [85]. The latter is the most commonly used type of adjuvant and there are currently three β -lactamase inhibitors in clinical use which are clavulanic acid, tazobactam, and sulbactam [86].

β -lactamase inhibitors, such as CA, are the mechanism used most successfully to combat antibiotic resistance [87]. CA is a β -lactamase inhibitor, very similar in structure to penicillins, known as a 'suicide inhibitor' whereby it permanently binds to the β -lactamase preventing the enzyme from hydrolysing the clavam ring of the penicillin by sacrificing itself as it does so. It has activity against both type 2 and 3 β -lactamases including those produced by *Staphylococcus spp* [88]. There are two types of clavam: (1) those with S-configuration containing sulphur in position one of the rings, these have strong antibacterial and anti-fungal properties however don't inhibit β -lactamase activity, and (2) the R-configuration where there is an oxygen molecule in position one instead of sulphur. CA belongs to this R-configuration group of clavams and was the first oxygen - analogue β -lactam to be isolated. The R-configured clavams have no anti-fungal properties, and are weak antibacterials but are good inhibitors of β -lactamases [89].

Clavulanic acid is co-formulated with amoxicillin as the proprietary GSK product Augmentin and now as the generic drug 'Co-Amoxiclav'. CA potentiates the activity of

amoxicillin by inhibiting β -lactamase activity, which results in amoxicillin being active against β -lactamase-mediated resistant strains of bacteria [88, 90]. Although Augmentin has been in use since the 1980s there have been very few cases of resistance arising in clinical isolates, this is incredibly important as there are few other adjuvant therapies so successful [85, 86]. The development of other β -lactamase inhibitors is also an area of significant industrial interest with a number in clinical development [91]. This is why research into the production and optimisation of antimicrobial potentiators is so important.

1.9 The GlaxoSmithKline *Streptomyces clavuligerus* lineage.

The industrial lineage of *S. clavuligerus* has undergone more than 30 years of random mutagenesis and selection for improved CA production at Smithkline Beecham Pharmaceuticals and subsequently GlaxoSmithKline. The initial patent culture of *S. clavuligerus*, ATCC 27064 (=DSM 738 = NRRL 3585) [47], was obtained by Beecham and re-streaked, selecting a single CA producing colony, and designated it *S. clavuligerus* SC2 (Figure 1.4). Chemical mutagenesis (unspecified in the historical Smithkline Beecham literature) was used to generate a branch point in the lineage, leading to *S. clavuligerus* SC3 and *S. clavuligerus* SC4 (Figure 1.4). The *S. clavuligerus* SC3 branch was not taken any further in the development process by Beecham and represents a dead end within the industrial lineage. *S. clavuligerus* SC4 subsequently underwent a number of rounds of mutagenesis/selection to increase CA titre, using chemical (again unspecified), UV and ionising radiation, prior to the selection of *S. clavuligerus* SC5 (Figure 1.4). Additional UV mutagenesis of *S. clavuligerus* SC5 gave rise to the stable strain *S. clavuligerus* SC6 for the production of CA (Figure 1.4). Morphologically these strains all undergo sporulation and produce the green-grey spore pigment characteristic of the species, though *S. clavuligerus* SC2 and SC6 exhibit the most consistent sporulation phenotypes (Figure 1.4). There is an increase in the CA titre in these strains as the lineage advances, with a steady increase in production from *S. clavuligerus* SC2 to *S. clavuligerus* SC6 (These data cannot be disclosed due to confidentiality).

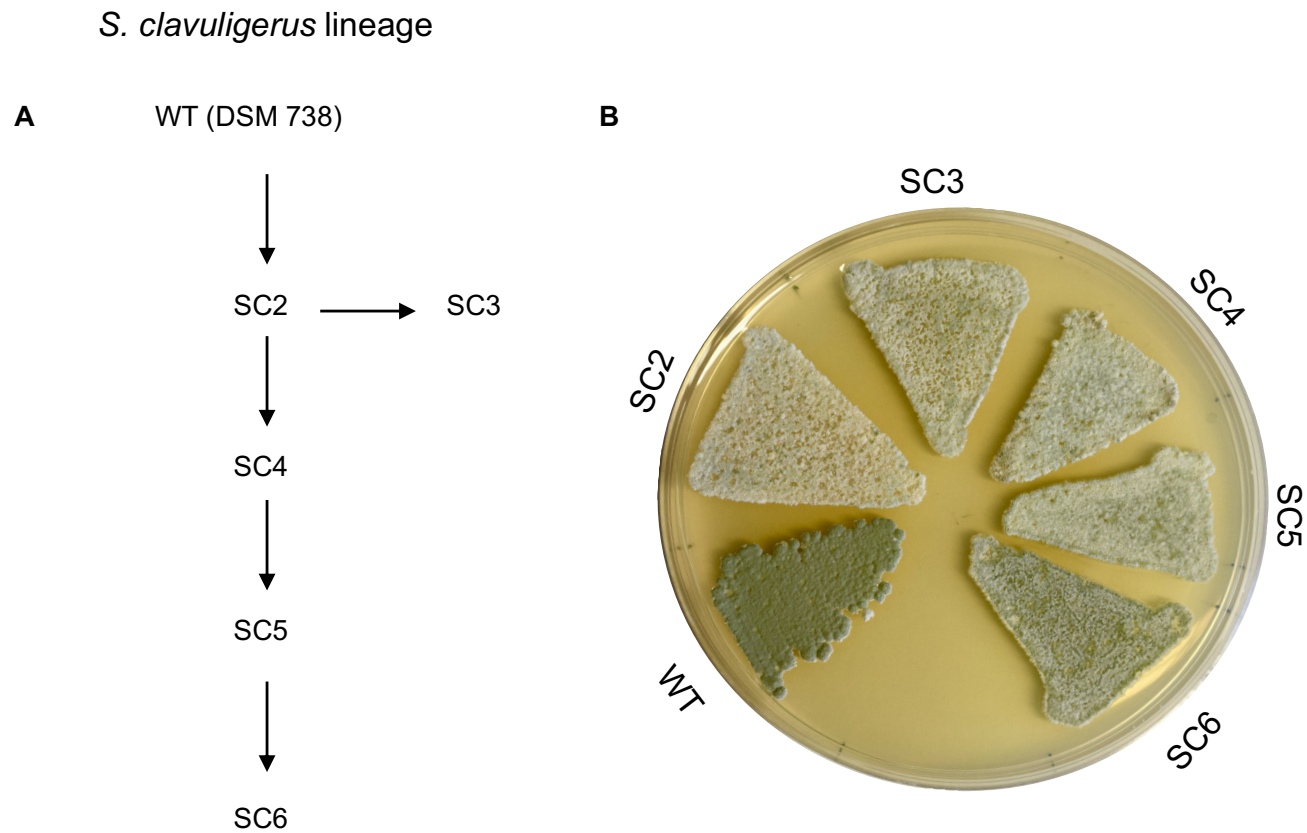


Figure 1.4 A. Representation of the first six strains in the *S. clavuligerus* industrial lineage and B. the morphologies of the strains are shown on the industrial media L3M9.

1.10 Microbial Metabolism

All organisms are comprised of chemicals and bacteria are no different. There are a collection of highly conserved chemicals, the intermediates and products of central metabolic pathways, which are shared by most bacteria [70]. 'Primary metabolites', the products of these central pathways amount to only several hundred compounds in most organisms. Traditionally, the separation of 'primary' and 'secondary' metabolism was widely adopted across biology, yet there has been an implication that secondary metabolism is less important than primary metabolism [41, 70, 92]. Generally primary metabolism refers to those pathways that normally give rise to single intermediate products, and the products of these pathways are utilised for 'usual' cellular functions [41, 70]. It is these pathways that supply the precursors for biosynthesis of specialised metabolites, where secondary metabolites are distinct products produced by specific taxonomic groups that are non-essential for the life of the organism [93]. Genetic manipulation can be used in a targeted approach to manipulate enzymes in the primary metabolism and increase the pools of precursor molecules in order to subsequently increase output of antibiotics [73].

1.11 Central node of metabolism

In *Streptomyces* the production of specialised metabolites such as antibiotics relies heavily on the availability of precursor molecules. These are often amino acids, lipids, or intermediate compounds from pathways such as glycolysis and the TCA cycle. Carbon metabolism and the availability of intermediate compounds from these pathways plays an important, but not fully understood, role in the production of secondary metabolites [49].

Carbon flux plays a significant role in the association between intermediates from the TCA cycle and their availability as precursor molecules for CA biosynthesis [49]. G3P is converted by GAPDH to 1-3-bisphosphoglycerate in glycolysis and then proceeds to the TCA cycle via pyruvate. As carbon flux directs approximately 80% of G3P through the glycolysis pathway, and only 20% to gluconeogenesis and CA production, it is unreasonable to think that merely increasing intracellular levels of G3P is a sustainable way to increase CA biosynthesis. Instead to increase CA production a better strategy is to manipulate the primary metabolism that directs precursor molecules towards the CA pathway [84].

There are 12 precursor metabolites upon which bacterial life depends, they are: glucose-6-phosphate, fructose-6-phosphate, ribose-5-phosphate, 3-phosphoglycerate, erythrose-4-phosphate, triose-phosphate, pyruvate, acetyl CoA, α -ketoglutarate, succinyl CoA, oxaloacetate, and phosphoenolpyruvate [94]. The regulation and production of these molecules is incredibly important and it had been reported that all 12 of the molecules can be synthesised via glycolysis, the pentose phosphate pathway, and the TCA cycle, which culminate at the PEP-PYR-OXA node of central carbon metabolism (Figure 1.5) [94]. Glycolysis and the pentose phosphate pathway are the main pathways by which carbohydrates are broken down, whereas the TCA cycle is required for both catabolism and anabolism. When an organism is grown on a gluconeogenic substrate that enters the central carbon metabolism via the TCA cycle, oxaloacetate or malate can be converted to pyruvate and PEP gluconeogenically to be utilised in the synthesis of glycolytic intermediates. There are numerous important enzymes that make up this anaplerotic node which essentially acts as a roundabout for carbon flux: accepting carbon from different pathways and directing it in the appropriate direction [95].

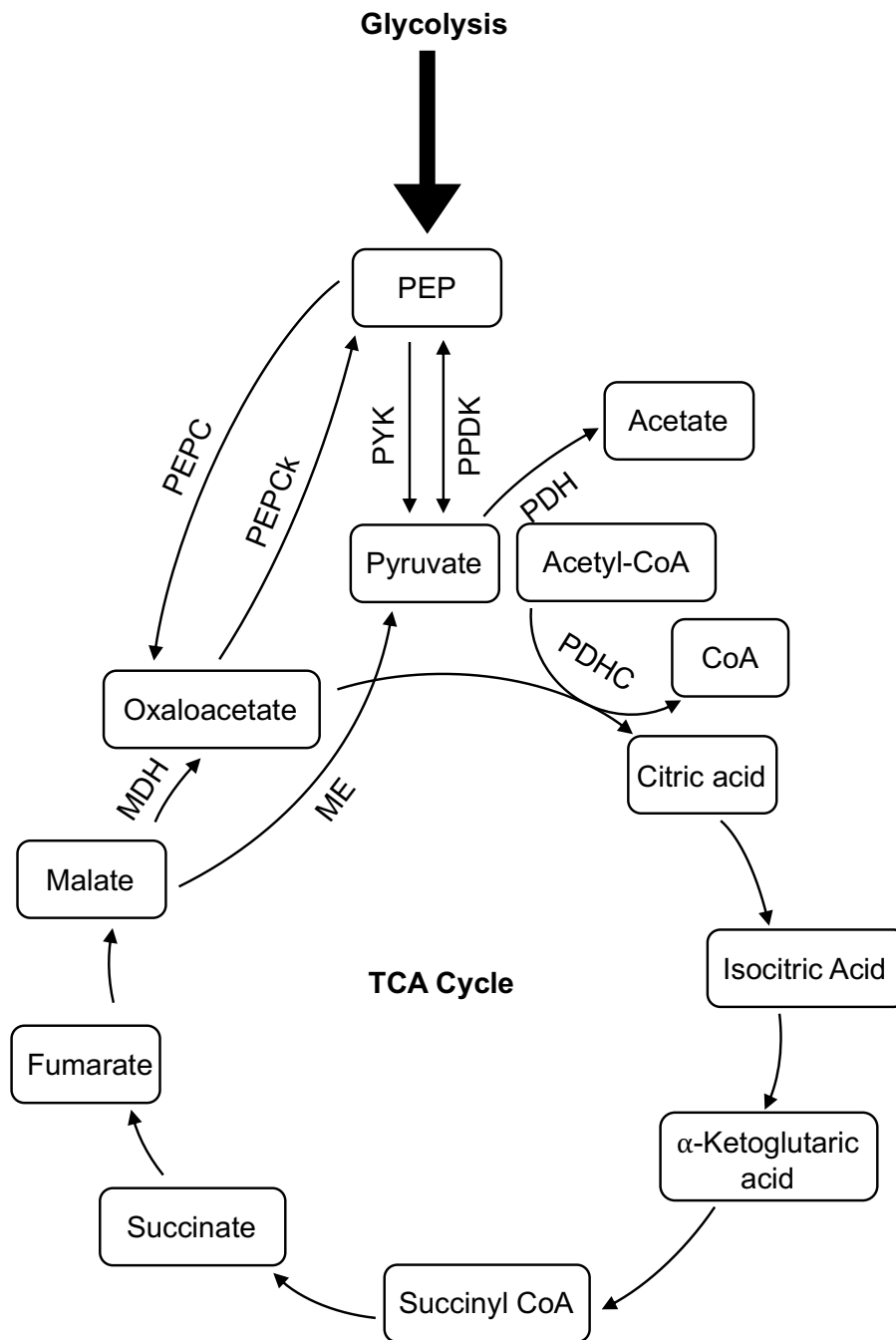
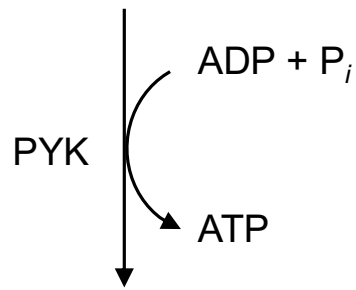


Figure 1.5 PEP-PYR-OXA node of central carbon metabolism. This node links glycolysis with the TCA cycle. Enzymes involved in the node include; pyruvate phosphate dikinase (PPDK), PEP carboxykinase (PEPCK), malate dehydrogenase (MDH), malic enzyme (ME), Phosphoenolpyruvate carboxylase (PEPC), Pyruvate carboxylase (Pyc), Pyruvate dehydrogenase (PDH), and Pyruvate dehydrogenase complex (PDHC).

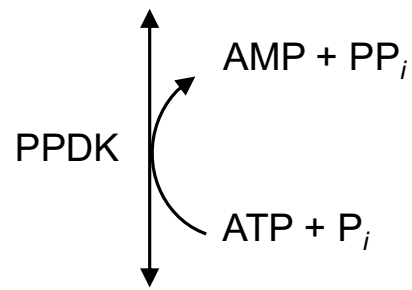
There are many enzymes that can be involved in this node and different organisms have different enzyme combinations (Figure 1.5). Some of these enzymes include Pyruvate kinase (PYK) which catalyses the glycolytic reaction of PEP to pyruvate, and the inverse reaction, pyruvate phosphate dikinase (PPDK), on which this thesis focuses. PPDK catalyses the reversible reaction of pyruvate to PEP in three partial reactions. PPDK although typically considered a gluconeogenic enzyme is likely to have originated as a glycolytic enzyme and has been shown to function glycolytically in bacteria, archaea, and protists. The functional direction of this enzyme can be influenced by environmental factors such as temperature and pH, a low pH favouring the pyruvate forming glycolytic direction and a high pH favouring the PEP forming gluconeogenic direction [96, 97]. Phosphoenolpyruvate synthase (Pps) is another enzyme used to interconvert pyruvate to PEP and similarly to pyruvate phosphate dikinase functions in three partial reactions and strongly favours the gluconeogenic direction (Figure 1.6). This enzyme is not present in *Streptomyces spp* but is present in actinobacteria that do not have a PPDK [97]. Three more enzymes involved in the node are PEP carboxykinase (PEPCK), malate dehydrogenase (MDH), and malic enzyme (ME), which together constitute the “malate shunt”: another pathway for PEP to be converted to pyruvate often used in organisms that do not possess a PYK [97, 98].

Phosphoenolpyruvate

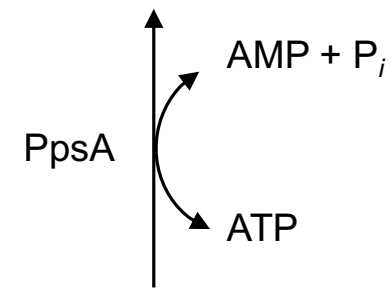
Pyruvate kinase



Pyruvate phosphate dikinase



Phosphoenolpyruvate synthase



Pyruvate

Figure 1.6 Three enzymes in central carbon metabolism that interconvert PEP and Pyruvate. Pyruvate kinase (PYK) a glycolytic enzyme that converts PEP to pyruvate. Pyruvate phosphate dikinase (PPDK) a bidirectional enzyme interconverting PEP and Pyruvate, Phosphoenolpyruvate synthase (PpsA) used to interconvert pyruvate to PEP and similarly to pyruvate phosphate dikinase functions in three partial reactions.

1.12 Pyruvate phosphate dikinase

In *Streptomyces* PPDK converts pyruvate, ATP and inorganic phosphate into AMP, PEP and diphosphate. The reaction is reversible in contrast to the pyruvate kinase (PYK) reaction and occurs in a three-step process (Figure 1.6). Previously PYK has been shown to be a good target for increasing antibiotic production [40]. There are two copies of *pyk* in the majority of *Streptomyces* and interestingly, there are also two copies of *ppdk*, indicating that there are two potential routes to converting pyruvate to PEP, one which is reversible and requires inorganic phosphate (PPDK) and one which is irreversible and utilises phosphate transfer from ATP (PYK). Moreover, it has recently been shown that at least one of the two paralogues of *ppdk* in *S. coelicolor* is upregulated 10-fold during the antibiotic production phase [99]. This potentially increases the robustness of function at this node of metabolism. Given that PPDK is reversible, this anaplerotic node could result in carbon flux being directed in either the glycolytic or gluconeogenic direction [100]. Carbon limitation can lead to a decrease in anaplerotic metabolism, resulting in a decrease in TCA cycle-derived biosynthesis [101]. It has been suggested that carbon flux through anaplerotic metabolism is responsible for a decrease in availability of the C3 precursor molecule pyruvate and consequently for limited production of CA [102]. Studies have shown that when growth medium is supplemented with amino acids such as arginine and threonine this decreased anaplerotic flux leads to an increase in the C3 precursor available for CA biosynthesis [103]. Targeted manipulation of enzymes in competing biosynthetic pathways including the TCA cycle has been shown to increase successfully the production of certain specialised metabolites [104]. Compared with random mutagenesis, a knowledge-based targeted approach to strain improvement can be used to make specific genomic changes to strains more accurately. This can be done by adding or removing genes and manipulating pathways to produce a desired

phenotype [78]. The genes encoding PPDK in *Streptomyces* offer a potentially interesting node for metabolic engineering.

1.13 Multiplicity of genes

A remarkable feature of *Streptomyces* genomes is that there are often multiple genes that encode the same putative enzyme function [27, 40]. It has been proposed that these redundant or contingent genetic loci are able to provide robustness in dynamic environments, enabling cells to cope with different physiological scenarios. The nature of these gene expansion events was traditionally considered to be the result of gene duplications. However it is becoming more apparent that horizontal gene transfer can play a role in acquiring this genetic multiplicity [40, 105, 106].

The phosphoenolpyruvate-pyruvate-oxaloacetate (PEP-PYR-OXA) node of metabolism is a major intersection of glycolysis, the TCA cycle, and gluconeogenesis, and represents a significant area of gene expansion in Streptomycetes [40]. The interconversion of primary metabolic intermediates at this node of metabolism represents a diverse range of expansion events, being highly variable and likely representative of the ecological diversity of these bacteria. The interconversion of metabolites at this point, linking catabolism and anabolism, highlights a major target for metabolic engineering in industrial organisms [107].

A key reaction at the PEP-PYR-OXA node is the interconversion of phosphoenolpyruvate and pyruvate [40, 70]. The two pyruvate kinases, converting PEP to pyruvate, were previously characterised [40]. However, it emerged during this work that the upregulation of PPDK activity on gluconeogenic substrates may play an important role in central carbon metabolism.

1.14 PPDK multiplicity

The PPDK enzymes were identified in the original gene expansion analysis by Schniete *et al*, [40]. The mean number of copies within the phylum was 1.36 copies of *ppdk* per species, however this masks the absence of these enzymes in many of the genera analysed by Schniete *et al*. Using these data, Table 1.1 shows the mean number of PPDK enzymes in different Actinobacterial species.

(data available: <https://doi.org/10.6084/m9.figshare.5772963.v1>)

Table 1.1 Mean number of copies of *ppdk* genes in different Actinobacterial genera

Genus	Mean number of <i>ppdk</i> genes
<i>Actinoplanes</i>	2
<i>Dactyosporangium</i>	3
<i>Micromonospora</i>	1.71
<i>Actinomycetospora</i>	2
<i>Lechevalieria</i>	2
<i>Streptomyces</i>	1.37
<i>Herbidospora</i>	2
<i>Streptosporangium</i>	1.5

Focussing on the genus *Streptomyces*, within the ActDES database [99] there are 287 representatives. One species lacks a *ppdk* homologue (*S. sp* NRRL WC3549), 197 species that have a single copy of *ppdk*, 78 species with two copies (gene expansion), 7 species that appear to have three copies of *ppdk*, one species with four copies (*S. bottropensis*), two species with five *ppdk* copies and a single species (*S. hygroscopicus* ssp. *Jinggangensis*) that has potentially eight *ppdk* copies within the genome. While a single copy of *ppdk* was present in 69% of *Streptomyces*, it would appear that gene expansion events have expanded this enzyme activity in some members of the genus. In the model species *S. coelicolor*, there are two copies of *ppdk* (SCO0208 and SCO2494) as well as in the industrially-exploited species, *S. clavuligerus*, (SCLAV1689 and SCLAV5441).

To further understand the phylogenetic pattern of distribution of *ppdk* in the *Actinomycetota* a phylogeny of the 612 PPDKs returned from a BLAST search of the ActDES database [108] was constructed in Figure 1.7. The NJ algorithm used by QuickTree [109] to build a phylogeny, showed a typical distribution of sequences, with clades largely within the order distribution of the *Actinomycetota* (Figure 1.7). The order *Streptomycetales*, exhibits a two phasic distribution of *ppdk* copies in the phylogenetic tree: a large clade representing the majority of the sequences in the ActDES database, and a second clade which represents the copies that have been expanded within *Streptomyces*. Given the difference in copy number and expansion frequency exhibited across the *Streptomyces*, this would suggest not a duplication event but that expansion has occurred by HGT, resulting in multiple enzymes with potentially adapted functions [40].

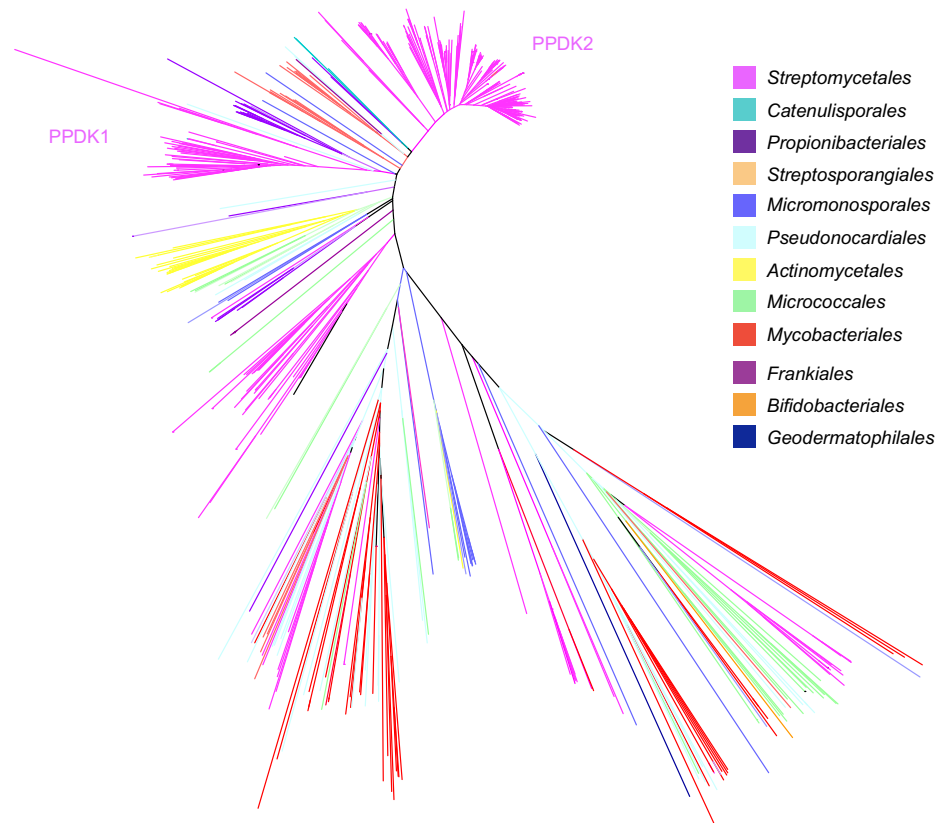


Figure 1.7 Pyruvate phosphate dikinase (PPDK) phylogeny. Built from amino acid sequences from all Actinobacterial species present in the ActDES database [108]. A total of 621 sequences were returned from a BLAST search, sequences were aligned using Muscle [110] and the tree was built using QuickTree using the NJ algorithm [109]

1.15 Aims:

- To understand the role of the PPDK enzymes in both model and industrial strains of *Streptomyces*
- To create PPDK knockdown and over expression mutants using CRISPRi technology
- To observe changes in phenotype, growth, and carbon utilisation between mutant and parental strains of *Streptomyces*
- To express, purify, and characterise three potentially important PPDK enzymes across model and industrial strains of *Streptomyces*

Chapter 2: Materials and methods

2.1 Bioinformatics

2.1.1 Multiple sequence alignments

Amino acid sequences used for multiple sequence alignment (MSA) were obtained from either StrepDB, NCBI, PDB or the ActDES databases [108]. Sequences were aligned using the Muscle alignment feature in Jalview [111].

2.1.2 Phylogenetic trees

Phylogenetic trees were constructed in two ways: in MegaX [112] by importing the Muscle alignment from Jalview and running the Maximum-likelihood algorithm.

In the case of large scale phylogenetic trees, amino acid sequences were extracted from the ActDES database [108] by BLASTP. The sequences were then aligned with MUSCLE V3.8.31(33), then alignments were scrutinised using Jalview V2.10.1 to ensure that at least 25 % coverage with the query was achieved. Sequences not meeting this threshold were discarded. Phylogenetic analysis of the alignments was conducted with QuickTree using the Neighbour-Joining algorithm [109] with trees visualised in FigTree 1v.4.4.

2.1.3 Predictions of protein structures

AlphaFold2 colab was used to predict structures of uncharacterised proteins using amino acid sequences from StrepDB as an input. PyMOL was then used to visualise and annotate the predicted structures PDB files for existing protein structures that were obtained from the RCSB Protein Data Bank.

2.2 Microbiology

2.2.1 Bacterial strains

Table 2.1. Bacterial strains used and generated in this project

Bacterial strain	Genotype	Reference
<i>E.coli</i>	Integrated pUZ8002 vector	Paget <i>et al.</i> , 1999 [113]
ET12567		
<i>E. coli</i> DH5 α		Taylor <i>et al.</i> , 1993 [114]
<i>E. coli</i> BL21		Invitrogen™
<i>E. coli</i> BTH101		Euromedex (ref. EUK001), Battestiet <i>et al.</i> , 2012
<i>S. clavuligerus</i>	Wild type	American Type Culture Collection
DSM 738		
<i>S. clavuligerus</i>	Integrated pGWS1370	This study
Δ control	vector	
<i>S. clavuligerus</i>	Integrated pIJ10257 vector	This study
+ control		
<i>S. clavuligerus</i>	Integrated pGWS1370	This study
Δ <i>ppdk</i>	vector Δ Sclav_1689	
<i>S. clavuligerus</i>	Integrated pIJ10257 vector	This study
+ <i>ppdk</i>	+Sclav_1689	
<i>S. clavuligerus</i>	GSK early production strain	GSK
SC2		

<i>S. clavuligerus</i>	Integrated pGWS1370	This study
SC2 Δ control	vector	
<i>S. clavuligerus</i>	Integrated pIJ10257 vector	This study
SC2 + control		
<i>S. clavuligerus</i>	Integrated pGWS1370	This study
SC2 Δ <i>ppdk</i>	vector Δ Sclav_1689	
<i>S. clavuligerus</i>	Integrated pIJ10257 vector	This study
SC2 + <i>ppdk</i>	+Sclav_1689	
<i>S. clavuligerus</i>	GSK late production strain	GSK
SC6		
<i>S. clavuligerus</i>	Integrated pGWS1370	This study
SC6 Δ control	vector	
<i>S. clavuligerus</i>	Integrated pIJ10257 vector	This study
SC6 + control		
<i>S. clavuligerus</i>	Integrated pGWS1370	This study
SC6 Δ <i>ppdk</i>	vector Δ Sclav1_689	
<i>S. clavuligerus</i>	Integrated pIJ10257 vector	This study
SC6 + <i>ppdk</i>	+Sclav_1689	
<i>S. coelicolor</i>	Wild type	Kieser <i>et al.</i> , 2000 [115]
M145		
<i>S. coelicolor</i>	Integrated pGWS1370	This study
M145 Δ control	vector	
<i>S. coelicolor</i>	Integrated pIJ10257 vector	This study
M145 + control		
<i>S. coelicolor</i>	Integrated pGWS1370	This study
M145 Δ <i>ppdk1</i>	vector Δ Sco_0208	

S. <i>coelicolor</i> M145 + <i>ppdk1</i>	Integrated plJ10257 vector +Sco_0208	This study
S. <i>coelicolor</i> M145 Δ <i>ppdk2</i>	Integrated pGWS1370 vector Δ Sco_2494	This study
S. <i>coelicolor</i> M145 + <i>ppdk2</i>	Integrated plJ10257 vector +Sco_2494	This study

2.2.2 Media

Table 2.2. Media and recipes used to grow bacterial strains in these experiments

Media name	Reagents
Autoinduction media ¹	2Y 929ml (see below) 50 x 5052 20 ml (see below) 20 x NPS 50 ml (see below) 1 M MgSO ₄ 1 ml
GYM ¹	CaCO ₃ 2g Glucose 4 g Malt extract 10g Yeast extract 4 g Agar 12 g
L3M9 ¹	Dextrin 0.3 g α - α -Trehalose 10 g K ₂ HPO ₄ .7H ₂ O 0.5 g NaCl 1 g MgSO ₄ 1 g CaCl ₂ 0.5 g Cassamino acids 2 g MOPS buffer 11 g Trace salts 1 ml ROKO agar 30 g Adjust pH to 6.8

Lysogeny Broth ¹	NaCl 10 g Tryptone 10 g Yeast extract 5 g Agar 16 g for solid media
M9 minimal media no carbon ^{1 2}	$(\text{NH}_4)_2 \text{SO}_4$ 1 g K_2HPO_4 0.5 g $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 0.2 g $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ 0.01 g Adjust pH to 7.0 – 7.2 Agar 10 g for solid media
Milk agar ¹	Milk powder 50g Tryptone 5g Yeast extract 2.5g Glucose 5g Agar 12.5g Adjust pH to 6.8-7.0
MS	Mannitol 16 g Soy flour 16 g Agar 16 g Make up to 1 l with tap water
YEME ¹	Yeast extract 3 g Malt extract 3 g

	Peptone 5 g
	Glucose 10 g
	Sucrose 340 g
2Y ¹	Tryptone 10 g
	Yeast extract 5 g
2xYT ¹	Tryptone 16 g
	Yeast extract 10 g
	NaCl 5 g
20 x NPS ¹	(NH) ₄ SO ₄ 66 g
	KH ₂ PO ₄ 136 g
	NaHPO ₄ 142 g
50 x 5052 ¹	Glycerol 250 g
	D-glucose 25 g
	α-lactose 100 g

¹ All media is made up to 1 L with dH₂O unless otherwise stated

² 166 mM of desired carbon source was added to the media

2.2.3 Buffers

Table 2.3. Buffer recipes used throughout this thesis

Buffer name	Reagents
Agar trace salts ¹	FeSO ₄ 10 g MnSO ₄ 10g ZnSO ₄ 10 g
Coomassie Blue Staining buffer ¹	0.1% Coomassie Brilliant Blue R-250 50% methanol 10% acetic acid
Destaining buffer ¹	40% methanol 10% acetic acid 50% dH ₂ O
IMAC Buffer A (pH 8.0) ¹	20 mM Na ₃ PO ₄ 500 mM NaCl 40 mM Imidazole
IMAC Buffer B (pH 8.0) ¹	20 mM Na ₃ PO ₄ 500 mM NaCl 500 mM Imidazole

Lamelli Buffer ¹	2% SDS 5% β-mercaptoethanol 10% glycerol 0.001% bromophenol blue 0.0625 M Tris-HCl pH 6.8
Nickel solution ¹	100 mM NiSO ₄
Stripping solution (pH 8.0) ¹	500 mM NaCl 100 mM EDTA
TAE Buffer (pH 8.3) ¹	40 mM Tris-HCl 20 mM glacial acetic acid 1 mM EDTA
TBS (Tris-Buffered-Saline) Buffer (pH 7.5) ¹	Tris HCl 7.88 g NaCl 8.76 g
TBS-T	99% TBS 1% Tween 20
TBS-T 5% milk	5% (w/v) Skimmed milk powder 95% TBS-T
Transfer buffer ¹	Tris base 3 g Glycine 14.4 g

SDS 1 g

MeOH 200 ml

¹ All buffers are made up to 1 L with dH₂O unless otherwise stated

2.2.4 Antibiotics

Table 2.4. Antibiotic stocks and working concentrations used throughout this thesis

Antibiotic	Stock concentration (mg/ml)	Working concentration (μ g/ml)
Ampicillin	100	100
Apramycin	50	50
Carbenicillin	100	100
Chloramphenicol	25 in EtOH	25
Hygromycin	50	50
Kanamycin	50	50
Naladixic acid	25	25

2.2.5 Culture conditions

S. clavuligerus strains were grown on L3M9 agar plates or in YEME liquid culture at 26°C for < 5 days. *S. coelicolor* and *S. venezuelae* strains were grown on MS and GYM agar plates respectively and TSB liquid culture at 30 °C for < 3 days. All *E. coli* strains were grown on LB agar plates or in LB broth at 37 °C for 12-16 hours.

2.2.6 Spore preparation

Streptomyces strains were grown on agar plates inoculated from a single colony for up to one week until a lawn of growth had emerged and sporulation had occurred. To make stocks, up to 5ml of sterile 20% glycerol was added to the plate and the spores agitated gently with a sterile cotton bud. The glycerol solution was then passed through a sterile syringe containing cotton wool to remove aerial hyphae. Stocks were aliquoted into cryotubes and stored at -20 °C [116].

2.2.7 Spore quantification

To quantify spores, a conical flask containing 50 ml of 2xYT plus antibiotics was inoculated with 1 ml of spores and incubated for 10 hours at 250 RPM and 26 °C or 30 °C for *S. clavuligerus* and *S. coelicolor* spores respectively. After 10 hours the OD₄₅₀ of 1 ml of culture was measured using a nanodrop spectrophotometer, the concentration of spores was calculated using the following equation:

$$C = (OD_{450} / 0.04) \times 4 \times 10^6 \text{ spores /ml} = OD_{450} \times 10^8 \text{ spores / ml [117]}$$

2.2.8 Growth curves using Cell Growth Quantifier (CGQ) [Aquila biolabs]

Spores (1×10^8) were pregerminated in 2 x YT media at 250 RPM at 26 °C and 30 °C for *S. clavuligerus* and *S. coelicolor* respectively. After 30 hours, cultures were spun at top speed in a Heraeus Megafuge 40R Centrifuge at 4 °C for 20 mins. The pellet was then washed in the same media as the growth curves were to be carried out in and this washing step was carried out for a total of three times. After the final spin, the supernatant was discarded and the pellet was resuspended in 50 ml of fresh media. The 50 ml culture was grown in a conical flask at 250 RPM at 26 °C or 30 °C for < 10 days until the culture reached stationary phase. The biomass of cultures was measured using the CGQ sensors. LEDs are used to shine light into the flask and photodiodes record light scattered back from the biomass in the flask. These data were received and measurements were recorded using the DOTS software.

2.2.9 Interspecies complementation experiments

Escherichia coli mutants were derived from the isogenic parental strain *E. coli* BW25113. The $\Delta pykAF$ double mutant was generated in the Hoskisson laboratory [40] and the Δpps mutant was obtained from the Keio Collection [118] and was kindly gifted by Dr Manuel Banzhaf (University of Birmingham). The *ppdk* containing vectors are listed in Table 5. Interspecies complementation generated *E. coli* mutants were derived from the isogenic parental strain *E. coli* BW25113. The $\Delta pykAF$ double mutant was created in the Hoskisson laboratory [40] and the Δpps mutant was obtained from the Keio Collection [118] and was kindly gifted by Dr Manuel Banzhaf (University of Birmingham). The *ppdk* containing vectors are listed in Table 2.5.

E. coli growth curves were carried out in 200 µl volumes in 96 well plates in a SYNERGY HT microplate reader (BioTek) at 37 °C with shaking (250 RPM). The initial inoculum was from an exponentially growing culture (1% v/v). Growth was followed spectrophotometrically at OD₆₀₀ for 16 hours. Strains were grown in LB, or M9 medium plus either glucose (1%), pyruvate (0.4%) or phosphoenolpyruvate (0.4%).

Growth of the cell was followed over time ($dN/dt = \mu N$; where N is the number of cells, t is time and μ is the specific growth rate in reciprocal units of time) and μ was calculated from the slope of a plot of the natural logarithm of biomass against time.

2.2.10 Cell dry weight

Cultures (50 ml) was spun at top speed in a Heraeus Megafuge 40R Centrifuge at 4 °C for 10 mins. After centrifugation, the supernatant was discarded and the pellet resuspended in 1 ml of growth medium and stored in a cryotube at -20 °C until use. To measure cell dry weight, micro glass fibre filter papers (from Fisherbrand, 1 µm pore size) were dried in the oven at 50 °C for ~ 3 hours and subsequently weighed. Biomass was applied to the dried filter paper and filtered through a Buchner Funnel (KIF Laboport and NALGENE 180 PVC metric tube) and subsequently washed 3 times with 1 ml dH₂O. After washing the biomass on the filter papers were returned to the oven and dried to a constant weight. The cell dry weight was determined by subtracting the weight of the dried filter paper alone from the weight of the dried filter paper and biomass.

2.2.11 Actinorhodin assay

Cultures (50 ml) were centrifuged in a Heraeus Megafuge 40R Centrifuge at 4200 x g and 4 °C for 10 mins. After centrifugation, 1 ml supernatant was taken and stored in

a cryotube at -80 °C until use. 500 µl of supernatant was added to an equal volume of 1 M NaOH and mixed by vortexing. This was subsequently transferred to a cuvette and OD₆₀₀ measured using a NanoDrop 2000c spectrophotometer. The concentration of actinorhodin was determined using the Beer-Lambert law;

$$A = \epsilon cl$$

Where;

A = absorbance

ϵ = molar absorption coefficient M⁻¹cm⁻¹

c = molar concentration M

l = optical path length cm

2.2.12 Screening *Streptomyces* for phenotypes

Parental and mutant strains were grown on a variety of media including, LB, Milk, and MS agars to screen for phenotypic changes including, growth, sporulation, and antibiotic production. Strains were streaked using a continuous method from a single colony and multiple strains were streaked on a single plate for comparative purposes. Plates were incubated for < 7 days at 26 °C or 30 °C for *S. clavuligerus* and *S. coelicolor* respectively.

2.2.13 Biolog

*PM1 MicroPlates*TM from Biolog were used to investigate the carbon utilisation of strains. L3M9 or MS agar plates were inoculated from a single colony and incubated for < 3 days at 26 °C or 30 °C for *S. clavuligerus* and *S. coelicolor* respectively. Once a lawn had grown, but prior to sporulation, biomass was harvested from the agar plate with a sterile cotton swab and the plates inoculated following the Biolog inoculation

protocol. Once the plates were inoculated, they were incubated in the OmniLog at 26 °C or 30 °C or *S. clavuligerus* and *S. coelicolor* respectively in the presence of Biolog Redox Dye Mix D (100x) to monitor growth for 96 hours.

2.3 Molecular biology

2.3.1 Plasmids

Table 2.5. Plasmids used and created throughout this project

Plasmid	Size (kb)	Features	Reference
pGEM-T Easy	3.0	ori ColE1, <i>amp^R</i> , MSC within β -galactosidase	Promega
CRISPR			
pGWS1370	10.5	ϕ C31, ori ColE1, <i>apm^R</i> , gRNA scaffold, dCas9	Zhang <i>et al.</i> , [119]
p Δ 1689	10.7	ϕ C31, ori ColE1, <i>apm^R</i> , gRNA scaffold, dCas9, Δ <i>sclav_1689</i>	This study
p Δ 0208	10.7	ϕ C31, ori ColE1, <i>apm^R</i> , gRNA scaffold, dCas9, Δ <i>sco_0208</i>	This study
p Δ 2494	10.7	ϕ C31, ori ColE1, <i>apm^R</i> , gRNA scaffold, dCas9, Δ <i>sco_2494</i>	This study
pIJ10257	6.4	oriT, <i>hyg^R</i>	Genscript
p+1689	9.1	oriT, <i>hyg^R</i> , <i>sclav_1689</i>	This study

p+0208	9.1	oriT, hyg ^R , <i>sco_0208</i>	This study
p+2494	9.1	oriT, hyg ^R , <i>sco_2494</i>	This study
<i>ppdk</i> over expression			
pET-14b	4.6	ori ColE1, <i>amp^R</i> , thrombin site, 6xHis	Genscript
pE1	7.4	ori ColE1, <i>amp^R</i> , thrombin site, 6xHis, <i>sclav_1689</i>	This study
pE2	7.4	ori ColE1, <i>amp^R</i> , thrombin site, 6xHis, <i>sco_2494</i>	This study
pE3	7.4	ori ColE1, <i>amp^R</i> , thrombin site, 6xHis, <i>sco_0208</i>	This study

2.3.2 Preparation of chemically competent *E. coli*

To make competent *E. coli* cells, 10 ml of LB was inoculated with a single *E. coli* colony and grown overnight at 37°C, 200 RPM. After 12-16 hours, the 10 ml overnight culture was added to 40 ml fresh LB and grown until it reached an OD₆₀₀ of 0.6. The culture was then spun down in a benchtop centrifuge at 2800 RPM, 4 °C for 10 mins. The supernatant was removed and the cells resuspended in 10 ml cold 0.1 M CaCl₂ and then left on ice for 20 mins. The cells were then centrifuged again at 2800 RPM, 4 °C for 10 mins, the supernatant discarded and the cells resuspended in 0.1M CaCl/15% Glycerol. Finally, cells were aliquoted into tubes and flash frozen in liquid nitrogen.

2.3.3 Plasmid miniprep

All plasmids were isolated from *E. coli* by inoculating 10 ml of LB (plus required antibiotics) with a single colony. Cultures were grown at 37 °C 250 RPM overnight and centrifuged for 10 mins at 2800 RPM., the supernatant was then discarded and the pellet used for extraction according to the Wizard[®] Plus SV Minipreps DNA Purification System protocol from Promega. Plasmid concentrations were determined using a Nandrop 2000c Spectrophotometer.

2.3.4 PCR clean up and Gel extraction

For PCR clean up and gel extraction, a 1:1 ratio of binding buffer to product was used before adding to the column. The Wizard[®] SV Gel and PCR Clean-Up System

[Promega] protocol was followed eluting the DNA in 20 µl of dH₂O and concentrations were determined using a Nanodrop 2000c Spectrophotometer.

2.3.5 Genomic DNA extraction

Genomic DNA (gDNA) was isolated from 5 ml cultures of *Streptomyces* grown for < 72 hours. Cultures were centrifuged at 10,000 x *g* for 10 mins and supernatant discarded. Pellets were resuspended in 300 µl of 50 mM/20 mM Tris-EDTA with 5 mg/ml and 0.1 mg/ml RNase. (111 µl of EDTA and 111 µl of Tris with 75 µl of 20 mg/ml Lysozyme + 3 µl RNase) were incubated for 30 mins at 37 °C. 50 µl of 10% sterile SDS was added and mixed thoroughly before adding 85 µl 1 M NaCl. After being transferred to phase lock Eppendorf tubes, 400 µl phenol/ chloroform/ isoamyl alcohol (25:24:1) was added, the reaction was vortexed for 30 secs and centrifuged at 10,000 x *g* for 10 mins. The aqueous phase was transferred to a new tube, and the chloroform step repeated three times. The aqueous phase was transferred to a new tube and 0.5 ml isopropanol added This was mixed by inverting and incubated at room temperature for 5 mins. After 5 mins this was spun down for 10 mins at 10,000 x *g* and the supernatant removed to leave the pellet. The pellet was then washed with 1 ml cold 70% ethanol before being poured off and the pellet left to dry. Finally, the DNA pellet was re-suspended in 20 µl sterile nuclease-free water.

2.3.6 RNA extraction

Strains for RNA extraction were grown in liquid culture (2xYT) for 48 hours at 250 RPM at 26 °C or 30 °C for *S. clavuligerus* and *S. coelicolor* respectively. After the 48 hours 1ml of culture was collected and was centrifuged at 10,000 x *g* for 5 mins [Beckman Coulter, Microfuge 20]. The pellet was then resuspended in RNA and stored

at -80 °C until required. RNA extraction was carried out using the RNEasy kit from QiaGen and using the Actinobase adapted protocol for RNA extraction from actinobacteria [116]. RNA was quantified using the NanoDrop 2000c spectrophotometer and either used immediately or stored at -80 °C until required.

2.3.7 PCR reaction for sgRNA amplification

sgRNA amplification was performed using CRISPR primers from Table 2.6, GoTaq PCR master mix by Promega, and pGWS1370 plasmid DNA as the template. The PCR conditions were as follows: initial denaturation 95 °C for 2 mins, further denaturation at 95 °C for 1 min, annealing at 42 °C-65°C for 1 min, extension at 72 °C for one min, this was repeated for 29 cycles. A final extension of 72 °C for five mins was carried out then the reaction was stored at 4 °C.

2.3.8 Colony PCR for *E. coli*

Colony PCRs to screen clones were carried out by inoculating 25 µl of REDTaq® ReadyMix™ PCR Reaction Mix [Sigma-Aldrich] containing 10 µM of forward and reverse primer with a single colony and running the PCR according to the manufacturer's guidelines.

2.3.9 RT-PCR

RT-PCRs were carried out using the OneStep *Ahead* RT-PCR Kit by Qiagen according to the manufacturer's guidelines. RNA that had been stored at -80 °C was thawed on ice and 10 ng of RNA was used as template for the RT-PCR, and 10 µM of both forward and reverse primer was used. The RT-PCR conditions were as follows:

reverse transcription 50 °C for 30 mins and initial activation 95 ° C for 15 mins, denaturation 94°for 30 seconds, annealing at 60 °C for 30 seconds, and extension at 72 °C for 1 min. Denaturation, annealing and extension steps were repeated for 30 cycles before a final extension of 72 °C for 10 mins.

2.3.10 Oligonucleotide primers

Table 2.6. List of oligonucleotide primers used in this work

Primer name	Sequence 5` - 3`
CRISPR	
SCLAV_1689 F	catgccatggaaactcgtaaggcggtcaagtttagagctagaaatag
SCO_0208 F	catgccatggcagaggcccttgctcggaagtttagagctagaaatag
SCO_2494 F	catgccatggaagccacgatgggcaagaaggtttagagctagaaatag
dCas9 R	ctagggatccaaaaaacccctcaagaccggttagaggccccaagg ggttatgctagtacgcctacgtaaaaaaagcaccgactcggtgcc
1689 Spacer R	aaactcgtaaggcggtcaa
0208 Spacer R	cagaggcccttgctcgaggaa
2494 Spacer R	aagccacgatgggacagaag
Spacer F	gctccagggggaaacgcct

2.3.11 Agarose gel electrophoresis

BioRad tanks were used for all agarose gel electrophoresis. Agarose gels were prepared by dissolving 1.5% (w/v) agarose in 1% TAE, and run in 1% TAE at 12.5 V/cm for one hour.

2.3.12 Restriction digest

To restriction clone insert DNA, either purified PCR product or plasmid, and backbone DNA, plasmid, was digested using 0.5 µl of each restriction enzyme [NEB] per ng of DNA as per the manufacturer's instructions and incubated at 37 °C for one hour. The digests were separated on a 2% agarose gel at 12.5 V/cm for one hour and the digested insert and backbone bands recovered using the Wizard® SV Gel and PCR Clean-Up kit by Promega.

2.3.13 Restriction cloning

Digested insert and backbone DNA fragments were ligated for one hour at room temperature or overnight at 4 °C using the T4 DNA ligase kit [NEB] and subsequently transformed into DH5α by heat shock. Transformations were plated on LB agar with selective antibiotics and incubated overnight at 37°C.

2.3.14 TA cloning

For subcloning purposes, pGEM-T easy was used. The T4 ligase kit was used to introduce DNA into this linearised plasmid with a vector to insert ratio of 3:1 and ligated at either room temperature for one hour or overnight at 4 °C. DH5α was transformed with the ligation reaction by heat shock and plated on selective media.

2.3.15 Transformation of *E. coli* by heat shock

To transform chemically competent *E. coli*, > 1 ng of plasmid DNA was added to a 50 µl aliquot of competent cells that had been thawed on ice. The cells were incubated on ice for a further 10 mins, then heat shocked at 42 °C for two mins before being returned to the ice. Cells were recovered in 950 µl of SOC broth for one hour at 37 °C. The exception to this was for competent *E. coli* ET12567/pUZ8002 cells which were recovered for three hours. The recovered cells were then plated onto selective media and incubated at 37 °C overnight.

2.3.16 Intergenic conjugation

The *E. coli* strain ET12567 containing the vector pUZ8002 was transformed with the oriT- containing vector to be delivered to *Streptomyces* and plated on LB agar and selected using chloramphenicol, kanamycin and the selective antibiotic for the plasmid. A single colony was then used to inoculate 10 ml of LB with the same antibiotics and was grown overnight at 37°C 200 RPM. After 12 – 16 hours the culture was diluted 1:20 with fresh LB and antibiotics and grown until it reached an OD₆₀₀ of 0.4-0.6. The cells were then collected by centrifugation at 2800 RPM, 4 °C for 10 mins and washed with fresh LB. This step was repeated twice and then cells were resuspended in 1ml fresh LB.

Approximately 1 plate of fresh spores were heat shocked at 50°C for 10 mins and allowed to cool. 500 µl of *E. coli* was added, mixed and pelleted. The supernatant was removed to leave a ~200 µl pellet. This was then plated on MS or L3M9 + 10 mM MgCl₂ and incubated at 30°C or 26°C respectively for 16 hours. After incubation the plates were overlayed with 1 ml dH₂O containing 0.5 mg nalidixic acid and the

selective antibiotics, and then incubated further. Exconjugants were then picked off and re-streaked on media containing selective antibiotics.

2.4 Protein expression

Pyruvate Phosphate dikinase proteins SCO0208, SCO2494, and SCLAV1689 from *S. coelicolor* and *S. clavuligerus* respectively were expressed in the *E. coli* strain BL21 from the pET14-b expression vector.

To express the proteins, 50 µl aliquots of One Shot™ BL21 Star™ (DE3) Chemically Competent *E. coli* were transformed with pET14-b containing each of the *E. coli* codon optimised genes by heat shock (refer to 2.5.5). The transformed BL21 cells were plated on LB agar containing 100 µg/ml Carbenicillin and incubated overnight at 37 °C. For each of the transformations, a single colony was picked and used to inoculate a seed culture of 10 ml LB with 100 µg/ml Carbenicillin which was incubated in a shaking incubator at 250 RPM and 37 °C until the cultures reached OD₆₀₀ 0.4–0.6. 50 ml of autoinduction media was inoculated with seed culture to a final concentration of 5%. Carbenicillin was also added to a final concentration of 100 µg/ml. This culture was then grown in a shaking incubator at 250 RPM and 26 °C for 18 hours. After 18 hours the cells were collected by centrifugation at 7,460 x *g* for 45 mins (JLA-9.100, Beckman Avanti JXN-26). Pellets were either used immediately or stored at -80 °C until required.

2.4.1 Cell lysis and preparation for purification

Prior to purification, the cells were lysed by sonication. If necessary, pellets were defrosted on ice and resuspended in 1.5 ml IMAC Buffer A (containing 10 mg/ml and 1mM DNase and PMSF respectively) for every 100 mg of pellet. Cells were sonicated

on ice at power setting 7 for 10 bursts of 15 secs with one min intervals. Lysed cells were centrifuged again at 14,970 x *g* for 30 mins (JA-25.50 Beckman Avanti JXN-26) to collect any unbroken cells. Clarified lysate was stored on ice and purification carried out immediately.

2.4.2 Immobilised Metal Affinity Chromatography (IMAC) .

A HisTrap™ FF Crude column [1ml; GE Healthcare] loaded with nickel, washed with dH₂O and equilibrated with 5 Column Volumes (CV) of IMAC Buffer A. The clarified cell lysate was suspended in IMAC Buffer A and loaded onto the HisTrap™ FF Crude 1 ml column [GE Healthcare] using the ÄKTA Pure FPLC platform [GE Healthcare]. Once sample loading was complete, the column was washed with Buffer A containing 40 mM Imidazole to remove any non-specifically bound proteins and leaving the His₆ tagged PPDK proteins bound to the column. All PPDKs were eluted by stepped elution over 10 Column Volumes (CV), the first 5 CV were eluted at 300 mM Imidazole and the remaining 5 CV at 500 mM. Eluant was collected in 1 ml fractions in a 96-well deep well plate. Proteins were verified by SDS-PAGE electrophoresis and total protein content was measured by Bradford assay [120].

2.4.3 SDS-Poly Acrylamide Gel Electrophoresis

Analysis of proteins was conducted by SDS-Poly Acrylamide Gel Electrophoresis (PAGE). Samples were mixed 1:1 with LDS sample buffer and heated at 85 °C for three minutes before loading 15 µl onto the gel. 5 µl of standard proteins, either Colour Prestained Protein Standard, Broad range (10-250kDa) [NEB] or Precision Plus Protein All Blue Prestained Protein Standards [Bio-Rad] were used. All samples were run on pre-cast NuPAGE™ 4 -12%, Bis-Tris, 1.0-1.5 mm, Mini Protein Gels [Invitrogen]

in MOPS buffer (pH 7.7) for < one hour at 25 V/cm constant using the Mini Gel Tank [Invitrogen].

2.4.4 PDK enzyme activity assay

To measure amount of PDK activity in samples of cell lysate and purified protein, a PDK activity assay kit [Abcam] was used. This colorimetric assay measures the conversion of the pyruvate substrate into an intermediate product which is detected by an enzyme mix and probe generating a stable product that can be read at 570 nm.

Chapter 3: Generation of *ppdk* mutants in *S. coelicolor* and *S. clavuligerus* using CRISPRi technology

3.1 Introduction:

S. clavuligerus is an important industrial organism used to produce clavulanic acid (CA) and, like many *Streptomyces* species, has been subject to much strain engineering to increase production of its desirable specialised metabolite. Historically a lot of strain development work, including some of the industrial strains used in this project, relied on random mutagenesis methods [78]. The scalability and rapidity of such methods offer advantages in an industrial setting, in that they generate initial wins of increasing titre. However, the lack of insight into and control of which genes are being altered can cause deleterious effects that become problematic later down the line. As a result, there has been a recent rise in the use of targeted approaches to increase specialised metabolite production in *Streptomyces*. It is well reported that precursor molecules for specialised metabolites are often derived from intermediates of primary metabolic pathways [27, 121]. This is exemplified by the biosynthetic pathway of CA, which begins with the condensation of L-arginine and glycolysis C₃ intermediate glyceraldehyde-3-phosphate as precursor molecules [122]. Therefore, the intersection between primary and secondary metabolism and the mechanisms controlling the flow of carbon through these pathways represents a powerful target for rationale-based strain improvement. There are several published reports highlighting the potential efficiency increases associated with engineering the central carbon metabolism in *Streptomyces*. Over-expressing or knocking down genes within the Embden-Myerhof (EM) pathway, TCA cycle, and altering flux around the anaplerotic node of central carbon metabolism can reduce bottlenecks and increase precursor availability for the synthesis of valuable specialised metabolites. This is demonstrated by a twofold increase in chloramphenicol production with *S. avermitilis* Δ *pfkA1* mutants by increasing the availability of the precursor D-erythrose-4-phosphate, an

intermediate from the pentose phosphate pathway crucial in chloramphenicol biosynthesis [40, 123].

Pyruvate phosphate dikinase, PPDK, makes up part of the anaplerotic node that links glycolysis and the TCA cycle (Figure 1.5). It catalyses the reversible reaction converting pyruvate to phosphoenol-pyruvate and is of importance because it facilitates gluconeogenic growth as well as glycolytic growth in organisms that possess this enzyme [94]. PPDK performs the opposite but reversible reaction to pyruvate kinase (Pyk), which was recently shown by Schniete *et al.*, [40] to be upregulated during antibiotic production when *Streptomyces* are grown on gluconeogenic substrates. PPDK is an important metabolic enzyme to target as it is one way of manipulating carbon flux in both the gluconeogenic and glycolytic directions. This can increase precursor availability in different areas of the TCA cycle, the pentose phosphate pathway, and glycolysis which is important for production of CA and other specialised metabolites.

CRISPRi technology dCas9 was used to generate the mutants described in this Chapter. This technology was chosen because it is non-invasive, represses gene expression at the transcriptional level, is highly specific, and is inducible in some strains [119]. CRISPR dCas9 is made up of two components, the spacer and guide RNA (sgRNA), and the catalytically de-active Cas9 endonuclease. The sgRNA has two parts: the protospacer, which is approximately 20 bp that is homologous to the target gene, and a Protospacer Adjacent Motif (PAM sequence) that directly precedes the protospacer. The Cas9 CRISPR machinery originates as a defence mechanism from *S. pyogenes* and the PAM sequence is the recognition site, here it is 'NGG' and the Cas9 protein cannot bind without it [124]. In the case of dCas9 it is an interference system which, when bound to the sgRNA, physically blocks transcription of the gene resulting in a lack of expression [125]. The CRISPRi

technology dCas9 was used to generate knockdown mutants to determine the role of PPDK in *Streptomyces*. The mutant strains were constructed in three strains of *S. clavuligerus* that represent different points in the industrial lineage of the current strain used by GSK for the commercial production of CA. The strain DSM 738 is the Wild-Type *S. clavuligerus* and is believed to be the progenitor of the GSK lineage and of the two other strains used in this work: *S. clavuligerus* GSK2 represents an early CA production strain, and *S. clavuligerus* GSK6 represents a later CA production strain. Both GSK2 and GSK6 are products of earlier random mutagenesis work carried out by GSK.

3.2 Aims and objectives of this Chapter:

The aims of this Chapter are to use CRISPR dCas9 to generate a collection of *ppdk* knockdown and *ppdk* overexpression mutants in four different *Streptomyces* strains: three industrial strains of *S. clavuligerus* (DSM738, GSK2, and GSK6) to investigate the effects on the industrial lineage, and the M145 strain of *S. coelicolor* as it is a model Streptomycete. Once generated, these mutants will form a foundational system for physiological study of the role of PPDK.

3.2.1 Specific objectives:

- Design guide RNA to knock down *ppdk* genes in both *S. clavuligerus* and *S. coelicolor*.
- Design plasmids to be synthesised for the over expression of *ppdk* genes in *S. clavuligerus* and *S. coelicolor*.

- Conjugate each of the four plasmids for the over expression and knock down of the two *S. coelicolor* *ppdk* genes (*SCO0208*, *SCO2494*) into the model strain *S. coelicolor* M145.
- Conjugate the plasmids for both the knockdown and over expression of *ppdk* (*SCLAV1689*) into the each of the three *S. clavuligerus* strains, DSM738, GSK2, and GSK6, which are the wildtype and two industrial strains respectively.

3.3 Designing guide RNA and generating CRISPR vectors

The dCas9 technology was used to generate knockdown mutants because, unlike other CRISPR systems, it does not result in a double stranded break in the DNA which can be disruptive to strains and potentially cause off-target effects [126].

The first step in constructing the CRISPR vectors was to identify an appropriate protospacer for the sgRNA. The protospacer was selected by identifying a region of ~20 bp homologous to the gene of interest, approximately 0.5 kb into the coding region of the gene that directly succeeds a PAM sequence. The sgRNA is where the CRISPR motif binds to the target DNA physically blocking transcription of the gene. This spacer was incorporated into a forward primer, and with the reverse primer used to amplify a fragment of the target gene approximately 200 bp in length.

The protospacers; AAACCTCGTCAAGGCGTTCAA, AAGCCACGATGGGACAGAAG, and CAGAGGCCCTTGCTCGAGAA were selected for the *ppdk* genes, SCLAV1689, SCO2494 and SCO0208 respectively. These were selected as they only appeared at this single point in the genome and were approximately 400-500 bp from the 5' end of the respective ORF blocking transcription of the full length CDS.

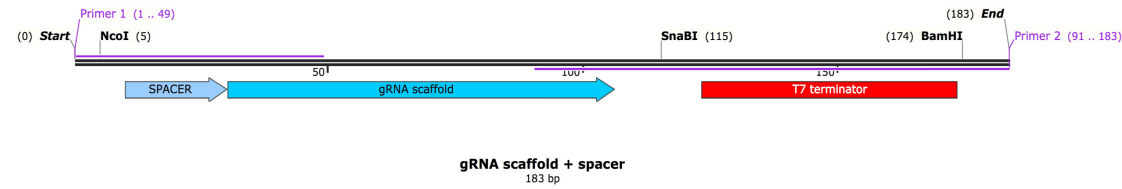
3.4 PCR to introduce 20 bp protospacer into sgRNA construct

The oligonucleotide primers SCLAV1689 F, SCO2494 F, SCO0208 F and dCas9 R, (Table 2.6) were used to simultaneously amplify the gRNA scaffold and incorporate a unique spacer for each target gene.

In order to find the optimal annealing temperature for these primers, the sgRNA was amplified via gradient PCR and the products analysed by agarose gel electrophoresis. Given the size of the spacer (20 bp) and the gRNA (78 bp), successful amplification

of the two in complex (Figure 3.1 A) should result in a band of ~100 bp. A band of the expected size was observed for SCLAV1689 and SCO2494 at all four annealing temperatures (Figure 3.1B), with the most intense band visible at 67 °C in both cases. In contrast, a 100 bp band was only observed at one annealing temperature (61 °C) for SCO0208. This indicates that the spacer and gRNA complex can be successfully amplified for each of the three target genes. These sgRNA complexes were then carried forward to construct the CRISPR vectors.

A



B

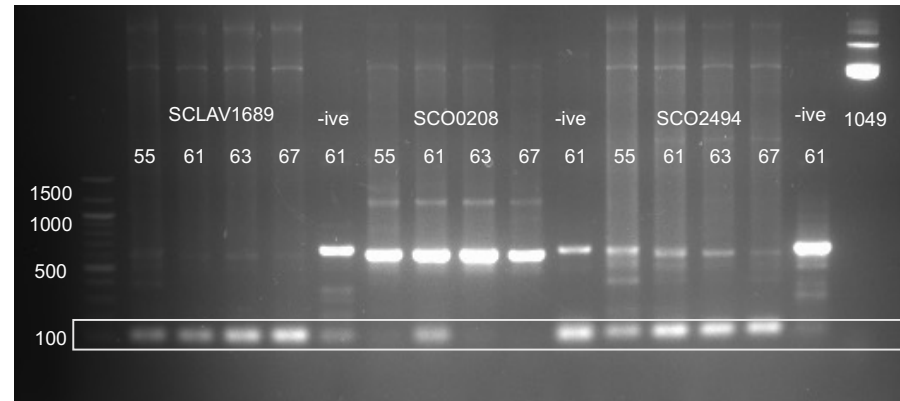


Figure 3.1 Amplification of spacer and CRISPR guide RNA. A Predicted sgRNA scaffold comprising of the spacer sequence, 20 bp, and the gRNA, 78 bp, amplified from CRISPR plasmid pGWS1049 DNA. B Amplification of the sgRNA for each *ppdk* by PCR which shows ~100 bp amplicons. 1049 represents a negative control vector containing the dCas9 machinery but no sgRNA. Ladder shown on the left is 100 bp hyperladder by Bioline.

The sgRNA complex PCR product was introduced into the pGEM-T easy vector by TA cloning, according to the manufacturer's instructions. This was to facilitate the subcloning of the sgRNA into the pGWS1370 CRISPR vector using the *Bam*HI and *Nco*I restriction sites present in the multiple cloning region of the pGEM-T easy vector. The ligation reaction was transformed into *E. coli* DH5 α via heat-shock as outlined in section 2.5.5 and subsequently spread on LB/Ampicillin/Xgal plates for selection and blue/white screening. The pGEM-T-Easy plasmid is compatible with blue/white screening, where blue colonies represent cells containing re-ligated vector and white colonies represent successful ligation of modified vector. White colonies were present on all transformation plates, but successful insertion of the sgRNA still had to be verified.

To confirm successful ligation of the sgRNA complex and pGEM-T Easy, vector plasmid DNA was extracted from overnight cultures of candidate colonies using a Wizard® *Plus* SV Minipreps DNA Purification kit. Each sample of plasmid DNA was used to perform two restriction digests, a single digest using *Bam*HI to linearise the vector to enable accurate interpretation of the size of the vector, and a double digest using both *Bam*HI and *Nco*I to extract the sgRNA from the pGEM-T Easy plasmid. In the event of successful insertion, the digest should give rise to a ~3000 bp fragment for the linearised pGEM-T easy vector, as well as a 98 bp fragment of DNA comprised of the gRNA sequence (78 bp) and the spacer (20 bp). When visualised via gel electrophoresis (Figure 3.2), the 1.5% agarose gel shows two bands in all digests. The first is a high intensity band beyond the limits of the ladder (>1500 bp) and the second is a much fainter band ~100 bp. The former band is likely pGEM-T while the latter band is consistent with the expected size of the sgRNA. The small size of the insert also explains the faintness of the observed band, as it comprises 3% of the total sequence length of the final vector and thus is only a small proportion of the DNA

loaded onto the gel. To the left of each of these double digests are single digests of each of the constructs to show the size of the linearised plasmids which is beyond the limits of the ladder. These verified pGEM-T-sgRNA vectors could then be used for subcloning into the final CRISPR vectors.

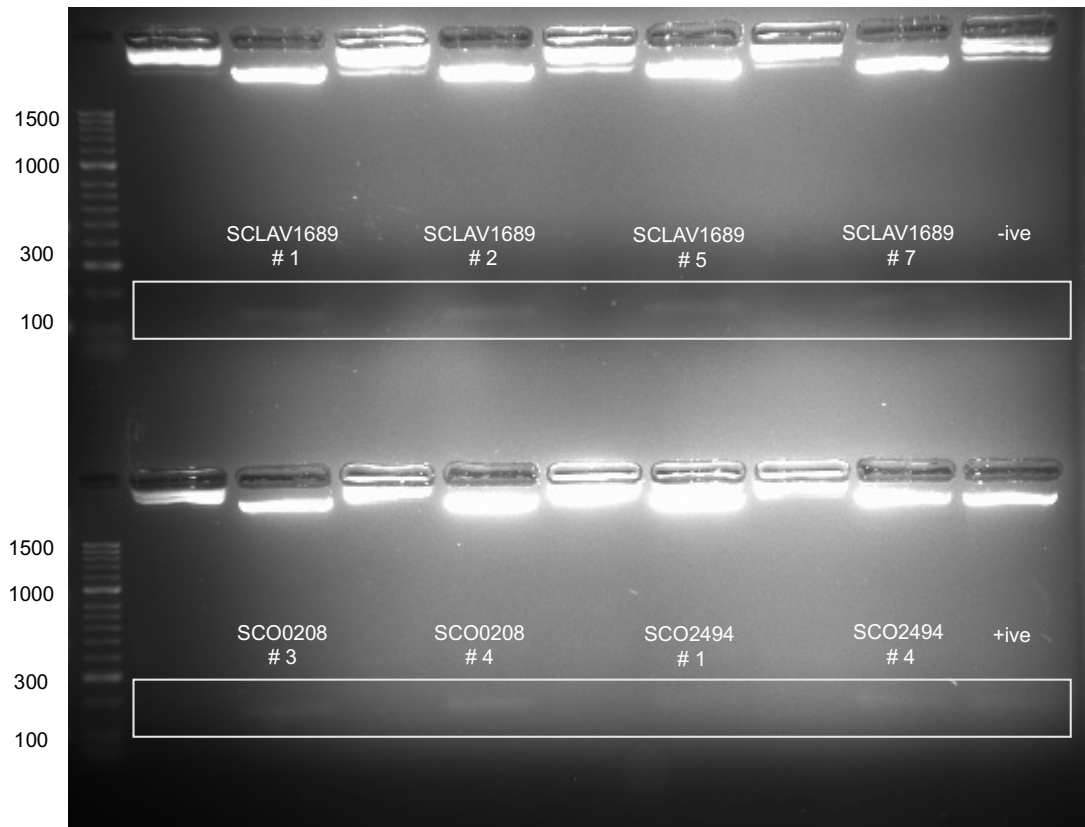


Figure 3.2 Confirmation of sgRNA in pGEM-T-Easy. Products of double restriction digest, 100 bp, next to the 1 kb plus ladder by NEB, this indicates that the gRNA + spacer had been excised from the pGEM-T Easy vector. To the left of each double digest is a single digest showing the size of the linearised vector with the insert. The negative control is pGEM-T-Easy with no insert and the positive control is the CRISPR vector digested showing the guideRNA excised ~78 bp, which should be ~20 bp smaller than the sgRNA fragments.

3.5 Cloning of sgRNA construct to generate final CRISPR vectors

Once the sgRNA inserts had been confirmed by digest and analysed by agarose gel electrophoresis, the CRISPR vector pGWS1370 was also digested using *Bam*HI and *Nco*I, and the two ligated. Ligation mixtures were transformed into *E. coli* DH5 α by heat shock and plated on LB/ apramycin (50 μ g/ml) plates and incubated at 37 °C overnight. Transformants were verified by sequencing in three independent clones to confirm the presence of the sgRNA prior to conjugation into *Streptomyces* (Figure 3.3).

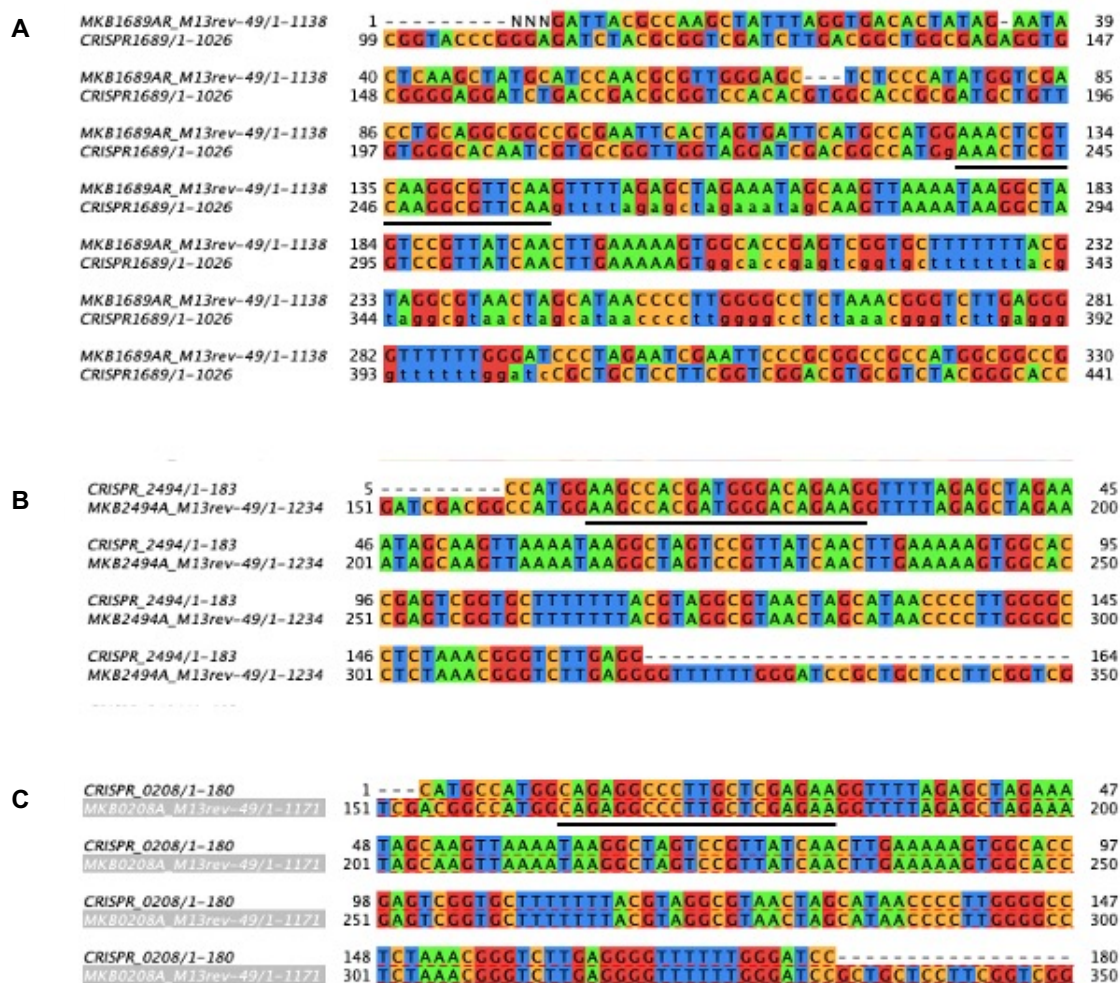


Figure 3.3 Confirmation of sgRNA in CRISPR vector A, B, and C show Sanger sequencing results of transformants containing the final CRISPR constructs for *SCLAV1689*, *SCO2494*, and *SCO0208* respectively aligned against the sequence of the final CRISPR vectors generated in silico using SnapGene. The results show the spacer for each gene is present in the relevant transformant which has been underlined.

3.6 Synthesis of *ppdk* over expression plasmids

While knocking down *ppdk* may result in interesting or useful phenotypes in the different *Streptomyces* strains, they will not conclusively demonstrate the role of PPDK. To verify the importance of PPDK in any observed phenotype, it is essential to carry out a corresponding set of complementation assays. Normally, complementation would be achieved by introducing a plasmid bearing the gene of interest into the mutant strain. However, the nature of CRISPR *dcas9* makes this approach impossible, due to the fact that the recognition site would be present on the plasmid copy of *ppdk* and thus recognised and silenced by the CRISPR machinery. Therefore, as an alternative approach, an extra copy of *ppdk* was introduced to parental strains. This was done using the pIJ10257 plasmid, produced by the Buttner group, which was modified from pMS81 to contain the *ermE** promotor, which is commonly used to introduce genes into *Streptomyces* [127]. This would enable comparison of selected phenotypes when *ppdk* expression is abolished (knockdowns), at normal levels (WT strains), or enhanced (overexpression) and thus help clarify the role of PPDK. Maps of the vectors for both the knockdown and overexpression mutants are shown in Figure 3.4.

Each of the *ppdk* genes for over expression were synthesised and blunt end cloned into the pIJ10257 expression plasmid by Genscript.

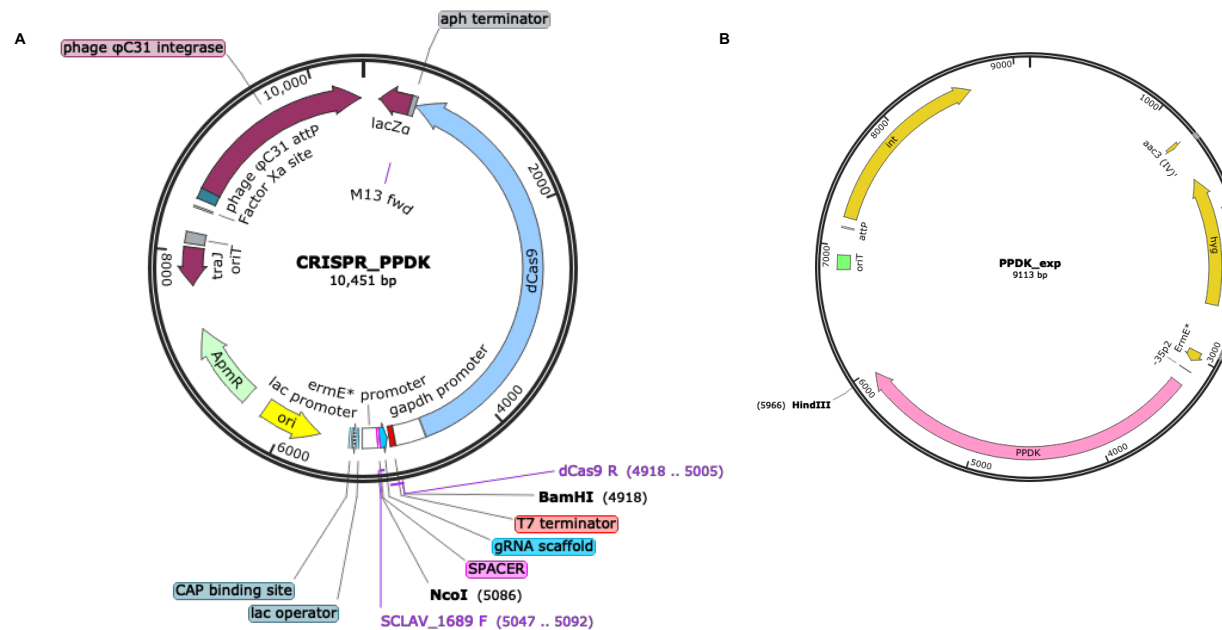


Figure 3.4 Vector maps of the CRISPR dCas9 knockdown and overexpression plasmids used in strain construction. A. CRISPR dCas9 vector used to generate knockdown mutants based on plasmid pGWS1370, contains dCas9 machinery downstream of sgRNA cloned into vector using *Bam*HI and *Nco*I restriction sites, downstream of *ermE** promoter. B. Vector used for overexpression of *ppdk* based on pIJ10257 plasmid, synthesised by genscript by blunt end cloning *ppdk* genes into *Nde*I and *Hind*III restriction sites, also under *ermE** promoter. In the cases of both CRISPR knockdown and overexpression vector 'PPDK' can represent any of the *ppdk* genes *SCLAV1689*, *SCO0208*, and *SCO2494*.

3.7 Conjugation of knockdown and overexpression plasmids into *Streptomyces* spp.

Plasmids to knockdown and overexpress *ppdk* SCLAV1689 were conjugated into *S. clavuligerus* DSM 738 and the GSK industrial strains *S. clavuligerus* GSK2 and GSK6 and plasmids to knock down and overexpress the two *ppdk* genes from *S. coelicolor* (SCO2494 and SCO0208) were conjugated into the wild-type M145 strain. They were conjugated from the non-methylating *E. coli* strain ET12567/pUZ8002 using one plate of dense spores per conjugation on the relevant medium. Exconjugants were picked after approximately two weeks and fresh plates containing apramycin (50 µg/ml) were inoculated from these single colonies.

As a negative control it was attempted to conjugate empty CRISPR vectors into the strains that did not contain any spacer. This meant theoretically that the dCas9 machinery could bind anywhere, which proved lethal for the *S. clauligerus* strains, but was successful in the model *S. coelicolor* strain M145. Once the knock downs and overexpression plasmids had been conjugated into their corresponding parental strains this resulted in a total of 15 strains (Table 3.1) which could be used to study the role of PPDK in *Streptomyces*.

Table 3.1 Mutants generated in each of the four different genetic backgrounds. Each tick represents an individual mutant strain.

Background strain	Δ SCLAV1689	+SCLAV1689	Δ SCO0208	+SCO0208	Δ SCO2494	+SCO2494	Δ control (pGWS1370)	+control (pIJ10257)
<i>S. coelicolor</i> M145			✓	✓	✓	✓	✓	✓
<i>S.clavuligerus</i> DSM 738	✓	✓						✓
<i>S.clavuligerus</i> GSK2	✓	✓						✓
<i>S.clavuligerus</i> GSK6	✓	✓						✓

3.8 RT-PCR used to confirm knockdown of *ppdk* genes

To confirm the knockdown of the *ppdk* genes, RT-PCR was carried out to measure relative expression of *ppdk* compared with the housekeeping gene *hrdB*. Successful knockdowns would be expected to have no, or very minimal, expression of *ppdk* compared to *hrdB* or to the parental strains. In contrast, unsuccessful knockdowns would have much more comparable levels of expression. To test this, RNA was extracted from each mutant and parental strain using the protocol for 'RNA extraction from actinobacteria' from ActinoBase [116], after 48 hours growth in 2xYT liquid media at 250 RPM and either 26 °C or 30 °C for *S. clavuligerus* and *S. coelicolor* strains respectively. 20 ng of RNA was used as the template for each RT-PCR, and the resultant products were analysed by agarose gel electrophoresis (Figure 3.5). Amplification of *hrdB* was used as a control to compare with the band of the *ppdk*. The results for the *S. clavuligerus* strains DSM 738 and GSK6 were inconclusive as on the agarose gel there are no visible bands amplifying either *hrdB* or *ppdk* from the RNA of the knockdown mutants (Figure 3.5A). If there was no band present in the lane for *ppdk* this would suggest it had been knocked down however there is no a band visible for *hrdB* either and this should still be present, meaning we cannot draw a conclusion from this result. However, in the *S. coelicolor* knockdown strains and the *S. clavuligerus* GSK2 $\Delta ppdk$ strain, it can be observed that expression in both parental and knockdown strains is different as illustrated by the brightness of the bands. The agarose gel shows the presence of an amplicon of the expected size, equating to *ppdk* in both the parental and knock down strains. However, in the knockdown strains ($\Delta ppdk$), the band is much lower in intensity when compared to *hrdB* than in the parental strains (Figure 3.5 B). This suggests that transcription is inhibited by the CRISPR dCas9 machinery and the genes have been successfully knocked down.

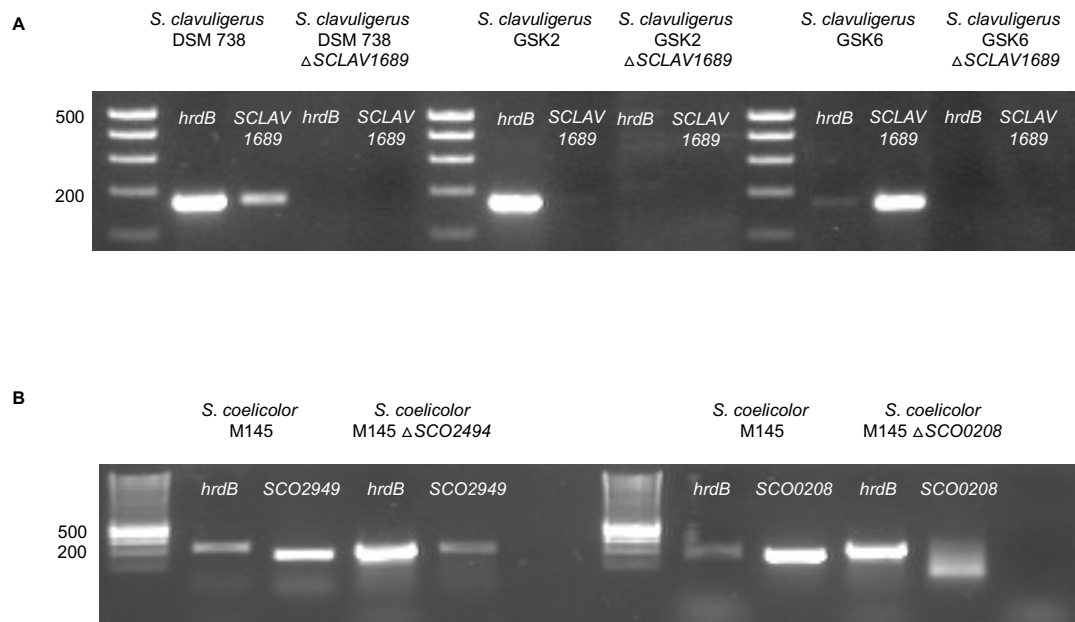


Figure 3.5 Testing of CRISPR dCas9 knockdowns using RT-PCR. A. RT-PCR of *S. clavuligerus* strains DSM 738, GSK2, and GSK6 and their respective *ppdk* knock down mutant from left to right, amplifying ~200 bp fragments of *hrdB* and SCLAV1689. DSM 738 and GSK6 results are inconclusive as neither *hrdB* or *ppdk* is shown to be amplified from the knockdown strains. GSK2 results show amplification of *hrdB* and SCLAV1689 from the parental GSK2 strains but only amplification of *hrdB* fragment from the knockdown mutant. B. shows *S. coelicolor* M145 and its respective knockdown mutants SCO0208 and SCO2494. In both cases *hrdB* and either SCO0208 or SCO2494 are amplified from both the Wild-Type and knockdown strain however the bands representing SCO0208 and SCO2494 are less bright in the knockdown strains. Each lane was loaded with 10 μ l of PCR product and the ladder is the 100 bp DNA ladder by New England Biolabs.

3.9 Discussion:

CRISPR genome editing was an exciting advancement in the field of *Streptomyces* research as compared with other model organisms such as *E. coli* or *S. cerevisiae* there are fewer tools and genetic manipulation is a lot more challenging [125]. In 2015 Huang *et al.* demonstrated that the Cas9 system could be used to delete the ACT and RED biosynthetic gene clusters in *S. coelicolor* to varying degrees of success [128]. In more recent years CRISPRi technologies such as CRISPR dCas9 have been more widely-adopted and used successfully to alter the expression of genes in important microorganisms such as Streptomyces which can be exploited for the purposes of natural product biosynthesis [129]. The CRISPRi technology has been used as recently as 2019 in *E. coli* to direct carbon flux to increase isopentenol biosynthesis, by knocking down genes controlling a competing biosynthetic pathway [130]. Despite this however, there are no published studies in the current literature documenting the use of dCas9 in *S. clavuligerus* specifically or in other industrial *Streptomyces* species. *Bacillus licheniformis* is one industrially relevant strain that has been manipulated using CRISPR Cas9 to increase the production of the valuable chemical 2,3-butanediol (BDO) by targeting the butanediol dehydrogenase enzyme [131].

There are several potential reasons why there are no published studies outlining the use of CRISPR technology in industrial Streptomyces like *S. clavuligerus*. The first being that there are still concerns about off target effects that result in cell toxicity as well as alter cell morphology [132]. Although CRISPRi technology is a more recent branch of CRISPR technology thought to reduced off target effects, there have also been some instances where it has been difficult to identify a suitable spacer that is adjacent to a PAM sequence. In this case the PAM sequence originating from *S. pyogenes* was NGG [133]. This meant that, due to the high GC content of the strains

used in this study, compatible PAM sequences occurred regularly throughout the genome. Thus, following a protocol derived from the work of Larson *et al.* and Qi *et al.* [124, 133], it was possible to design spacer sequences to target the three genes of interest and subsequently generated the mutant strains listed in Table 3.1.

In all, 15 strains were produced in this Chapter including controls. The mutants generated will facilitate the investigation of the role of PPDK in *Streptomyces* specifically. This seeks to help fill the knowledge gap surrounding this enzyme as although there has been research published regarding primary metabolic enzymes and the PEP-PYR-OXA node of central metabolism in bacteria, in general PPDK is often overlooked [94].

**Chapter 4: Changing the expression
of *ppdk* alters the physiology and
growth characteristics of
S. coelicolor and *S. clavuligerus***

4.1 Introduction

In this Chapter the physiology, carbon utilisation, and antibiotic production of the *ppdk* mutant *Streptomyces* strains, produced in Chapter 3, are investigated.

Whilst the model *S. coelicolor* is generally thought to have a wide substrate utilisation, this is not true for its industrial counterpart. *S. clavuligerus* is typically cultured in media containing glycerol or oil as carbon source, which is particularly popular in industrial fermentations due to the high titres of clavulanic acid (CA) produced [134, 135]. Simple sugars, such as glucose, represent more attractive sustainable options, however it has been reported numerous times that *S. clavuligerus* is incapable of utilising such sources. Despite this, the glucose permease and kinase genes required to process glucose are present in the *S. clavuligerus* genome, suggesting that this functionality was once present, but has since been lost during the years of adaptation to glucose-poor industrial fermentation media [136]. Although glycerol is used by many companies it is a relatively expensive carbon source, and rising prices make it an unsustainable primary carbon source for the long-term. It is already known that random mutagenesis and adaptation to a specific nutrient environment can reduce substrate utilisation of an organism and this has been shown in *E. coli* [137]. Reducing the expression of *ppdk* enables the investigation of how simple sugars are utilised when there is a reduced capacity to utilise gluconeogenic sources.

Chapter 3 described the production of mutant strains of both model and industrial *Streptomyces* and how this aims to build a more targeted approach to develop the industrial strains.

In this Chapter a large scale analysis of the substrate utilisation of both the parental strains and their subsequent mutants is undertaken. This was carried out using the phenotypic microarray PM1 plates from Biolog which analyses carbon utilisation using a redox dye-based system patented by Biolog.

This draws attention to how the substrate utilisation of the industrial strains differs from the model, and how the substrate utilisation of industrial strains has decreased as they have become increasingly adapted to the industrial environment. The Biolog experiment highlights the effect that reduction or over expression of *ppdk* has on carbon utilisation in these *Streptomyces* spp.

From the data on carbon source utilisation by the strains obtained from the Biolog experiment, two carbon sources were chosen to carry forward to experiments on utilisation of single carbon sources.

As CCR plays such a big role in the utilisation of carbon sources, the aim was to investigate the effect of growing the strains on a sole carbon source. To investigate if this alone would impact growth and if altering the expression of *ppdk* would further impact growth on these carbon sources.

The utilisation of different carbon sources also affects the availability of precursor metabolites for antibiotic production. If *ppdk* affects the utilisation of carbon and growth, this could lead to changes in the production of antibiotics by these organisms which is reflected in antibiotic titres.

4.2 Aims and objectives of this Chapter:

The aim of this Chapter is to utilise the mutant strains created in Chapter 3 to understand if the *ppdk* genes (*SCO0208*, *SCO2494*, and *SCLAV1689*) affect the carbon utilisation, growth rate, and/or antibiotic production of the *Streptomyces* strains and, if so, gain insight into the underlying mechanism.

4.2.1 Specific objectives:

- To investigate if altering the expression of *ppdk* impacts the phenotypic characteristics of two *Streptomyces spp.*
- To show how substrate utilisation of the strains has changed due to adaptation to industrial conditions and investigate if expression of *ppdk* genes affects that.
- To observe how the strains grow on sole carbon sources and compare the parental and mutant strains.
- To observe antibiotic production of the strains when grown on sole carbon sources and investigate if *ppdk* plays a role in the antibiotic titre obtained.

4.3 Analysis of nutrient utilisation using phenotypic Biolog arrays shows decreasing substrate utilisation in industrial *Streptomyces* spp.

Phenotypic analysis of the strains was undertaken using the Biolog PM1 phenotypic arrays with the OmniLog plate reading system. The Biolog PM1 plates are in a 96-well format, with each plate containing 95 different carbon sources and 1 negative control. Carbon sources investigated are shown in Appendix 1. This enables the strains to be screened for growth on a variety of substrates including amino acids, lipids, simple sugars, polysorbates, and esters to identify the optimum carbon source for their cultivation.

Each of the plates was inoculated with one of the parental or mutant strains and incubated in the OmniLog plate reading system at 26 °C or 30 °C, for *S. clavuligerus* and *S. coelicolor* respectively, for 96 hours. During this incubation period growth was measured via the redox dye-based system patented by Biolog. The metabolic activity of the strains is measured by mixing the cell suspension with a tetrazolium dye which is reduced to a purple formazan as the strains metabolise the carbon sources in the wells of the plate. Reduction of the tetrazolium is irreversible and the resulting purple compound can be measured colorimetrically at 590 nm using the OmniLog by Biolog [138]. It was expected that the model *S. coelicolor* M145 would grow on a wide range of carbon sources compared with any of the *S. clavuligerus* strains because this model organism has not been adapted to a specific environment in the same way that the industrial strains have. As the industrial strains are increasingly adapted to industrial fermentation conditions, it was expected that the strains would be adapted to utilise these specific carbon sources and lose the ability to grow on others all together. It was also expected that knocking down *ppdk* may reduce substrate utilisation in both organisms as it should limit carbon flux around the anaplerotic node and would abolish one route for utilising gluconeogenic substrates, resulting in

sources such as glycerol being harder to utilise. The results acquired from these Biolog experiments are represented in a heatmap created using the ggplot package in R. The heatmaps (Figures 4.1 and 4.2) show 34 carbon sources focusing on the results obtained from growing the strains on polysorbates, esters, and sugars.

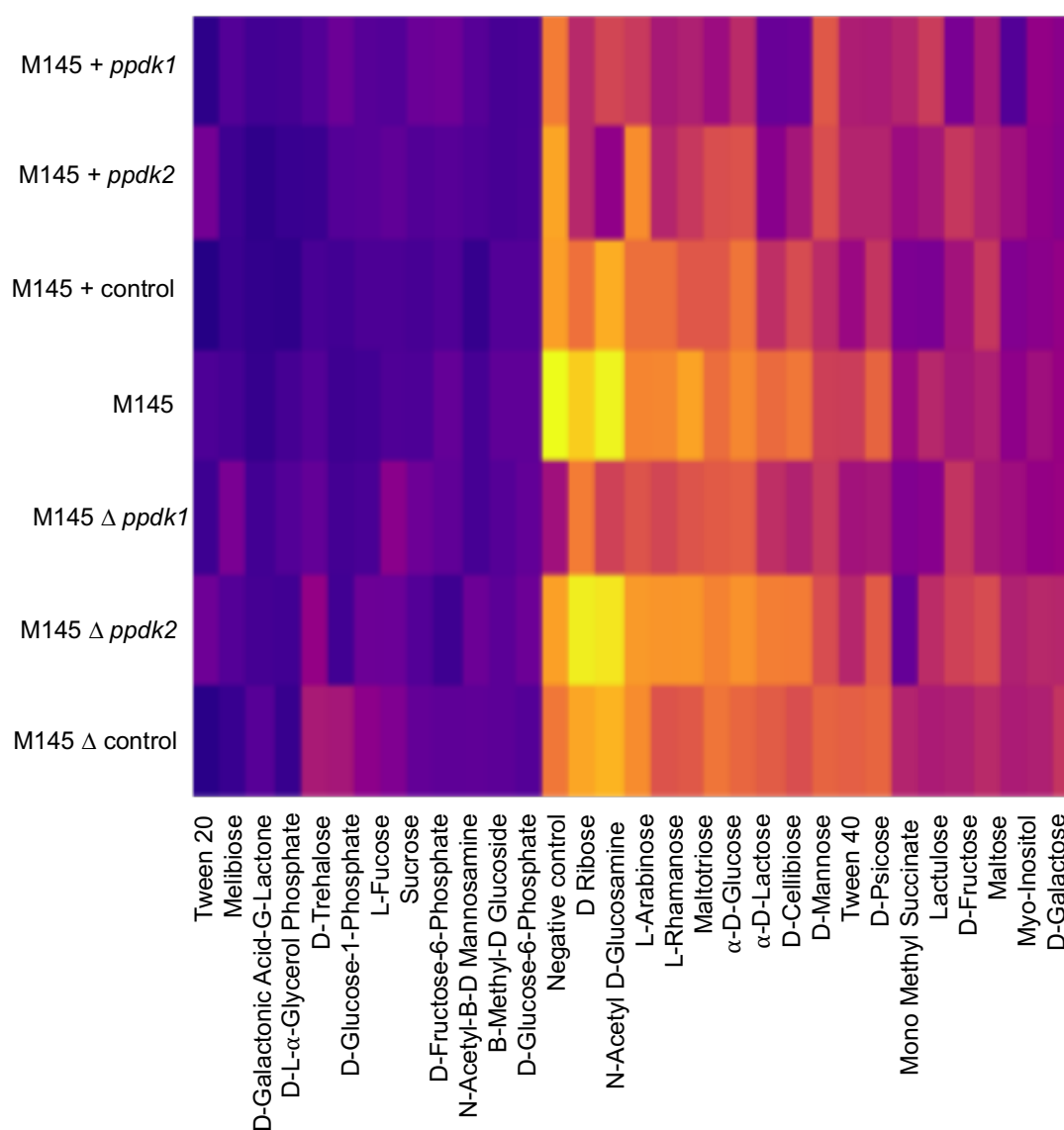


Figure 4.1 Heat map of growth *S. coelicolor* and its derivative strains on sugars, lipids, and esters from the PM1 Biolog plates. The growth of *S. coelicolor* and its derivative strains on 34 different carbon sources obtained by growing the strains on PM1 plates measuring the growth using the OmniLog by Biolog. Strains are listed on the y-axis and carbon sources on the x-axis. Negative control is no carbon source provided, purple represents no growth and yellow represents maximum growth.

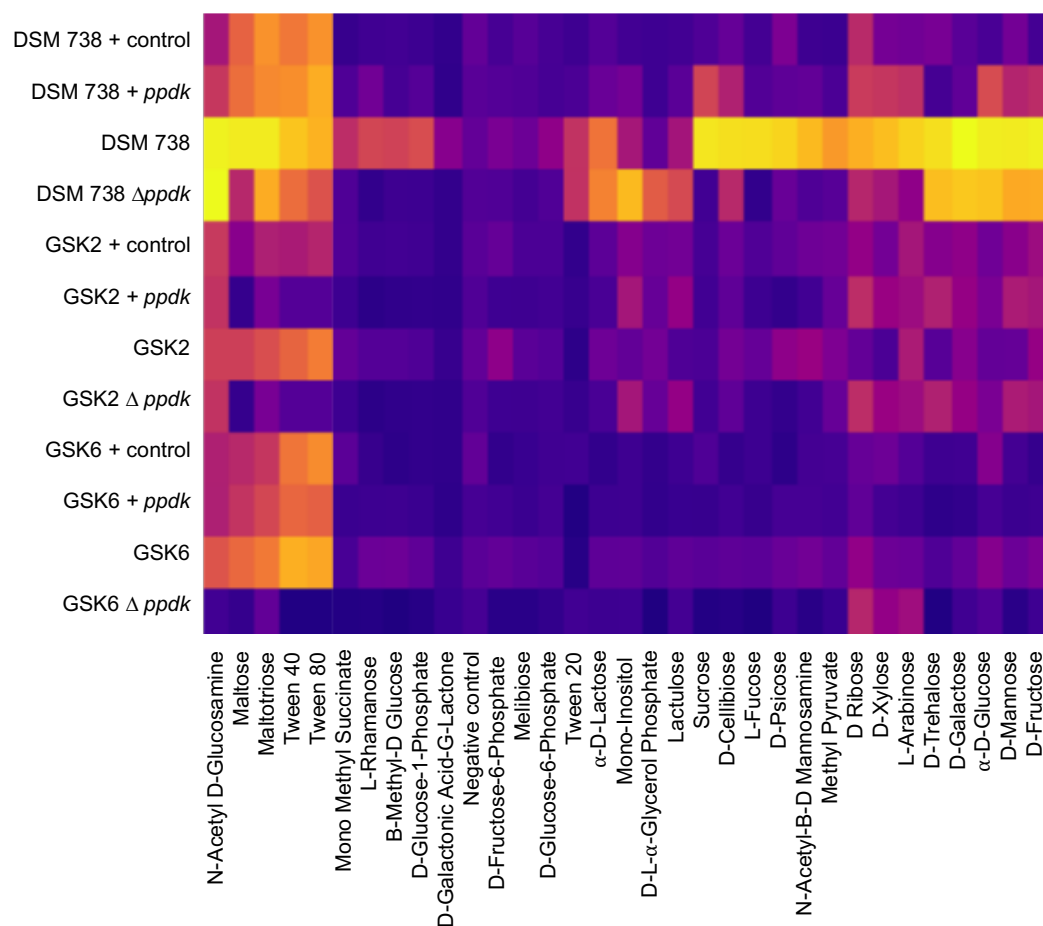


Figure 4.2 Heat map of growth *S. clauligerus* DSM738, GSK2, and GSK6 and their subsequent derivative strains on sugars, lipids, and esters from the PM1 Biolog plates. The heatmap illustrates the growth on 34 different carbon sources over 96 hours on the PM1 plates by Biolog. Strains are listed on the y-axis and carbon sources on the x-axis. The key shows purple represents no growth and yellow represents maximum growth. Heatmap groups carbon sources that illicit similar growth patterns.

The Biolog analysis focusses on how each parental strain and its mutants grow on sugars, polysorbates, and esters. As expected, the model *S. coelicolor* M145 grows on a wider range of carbon sources compared with any of the *S. clavuligerus* strains. The heat map (Figure 4.1) shows that the strains grow well on sugars such as glucose, lactose and maltotriose. However, the best carbon sources for growth were arabinose and ribose. The screen shows that the utilisation profiles of M145 and its mutants are broadly similar; however M145 $\Delta ppdk1$ (SCO0208) appears to be impacted most. This strain appears to show a reduced utilisation of some of these sugars, specifically the five carbon monosaccharides ribose and arabinose, which suggests that *ppdk1* plays an important role in utilising simple sugars. Interestingly this Biolog analysis also shows that M145 $\Delta ppdk2$ (SCO0208) has increased growth on these same sugars, suggesting that the two PPDK enzymes in *S. coelicolor* may play conflicting roles. Although knocking down either of the two *ppdk* genes in *S. coelicolor* impacts the capability to utilise some simple sugars, overexpressing either of the genes has little effect suggesting that overexpressing them is not relieving any bottlenecks in the system.

Although, as previously stated, *S. clavuligerus* is reportedly unable to utilise glucose despite encoding genes for the requisite cellular machinery [139] here the Biolog data shows that the strains are all able to grow on glucose as a single source to some extent, and the wild type strain and its knock down mutant grow well on glucose as a sole carbon source (Figure 4.2).

In *S. clavuligerus* DSM738, altering expression of the *ppdk* gene (SCLAV1689) by either knock down or overexpression, narrows the substrate utilisation of the strain. Although knocking down *ppdk* in this strain reduces the overall number of usable carbon sources (from 30 to 16), it does also appear to increase the overall amount of growth on some of the carbon sources, namely glycerol-1-phosphate and myo-inositol.

The early-derived industrial strain *S. clavuligerus* GSK2 and its derivatives (GSK2 $\Delta ppdk$ and GSK2 +*ppdk*) have increased growth on sugars such as maltose and maltotriose compared to the Wild-Type *S. clavuligerus* (DSM738) strain. Comparison of the late industrial strain *S. clavuligerus* GSK6 and its respective derivatives (GSK6 $\Delta ppdk$ and GSK6 +*ppdk*) shows that there is a clear reduction in substrate utilisation, with reduced catabolism of carbon sources compared with both *S. clavuligerus* DSM738 and *S. clavuligerus* GSK2. Yet there is a clear switch to utilisation of the lipid-like polysorbates such as Tween, which likely reflects the use of oil-based media in the industrial setting.

4.4 Investigating growth rates using single carbon sources.

Carbon utilisation analysis from the Biolog shows that *S. clavuligerus* DSM738 has the widest carbon utilisation profile of the industrial strains, decreasing with the first industrial strain *S. clavuligerus* (GSK2) and then decreasing further for the most recent industrial strain used, *S. clavuligerus* GSK6. As these strains have been adapted to industrial fermentation conditions, it is teleological that the strains have adapted to utilise these specific carbon sources and lose the ability to grow on others all together. The data also show that the mutants generated in Chapter 3, either knocked down or overexpressed *ppdk*, do not change the overall utilisation profiles of the strains. Instead, a change in the rate at which these carbon sources are utilised is observed.

To follow on from this large-scale analysis of carbon utilisation, two carbon sources were chosen for more detailed investigation into how the strains grow when on single carbon sources. The first carbon source chosen was glucose. Glucose is readily available as a waste product from hydrolysed bread and potato, making it highly sustainable and relatively cheap, which are desirable traits for industrial processes [6]. In addition, the ATP titre per molecule of glucose is high, making it a viable feedstock for industrial-scale antibiotic fermentations [7].

The second carbon source chosen was the polysorbate Tween 40. The strains all grow relatively well on this carbon source as it mimics the oil-based media currently used for the industrial production of clavulanic acid. When investigating the utilisation of single carbon sources, it was expected that the model organism *S. coelicolor* would grow well on glucose and its derivative M145 $\Delta ppdk2$ (SCO0208) was expected to grow even better on this sole carbon source, as the Biolog analysis showed increased utilisation of simple sugars when *ppdk2* (SCO0208) is knocked down. By comparison, it was expected that the *S. clavuligerus* strains would show increasing growth rate on

Tween as a sole carbon source as the industrial lineage progresses and that knocking down *ppdk* (*SCLAV1689*) would impede the growth rate due to the reduction in the ability to utilise gluconeogenic carbon sources.

To observe only the utilisation of these specific carbon sources, in each instance 166 mM of carbon was added to one litre of an adapted minimal media recipe [116]. This is to keep in line with previous unpublished studies on carbon utilisation in these industrial strains where they were grown on minimal media supplemented with 5 g/l of glucose (166 mM carbon). The strains were then grown on minimal media supplemented with either glucose or Tween 40 as the carbon source and their growth and antibiotic production measured.

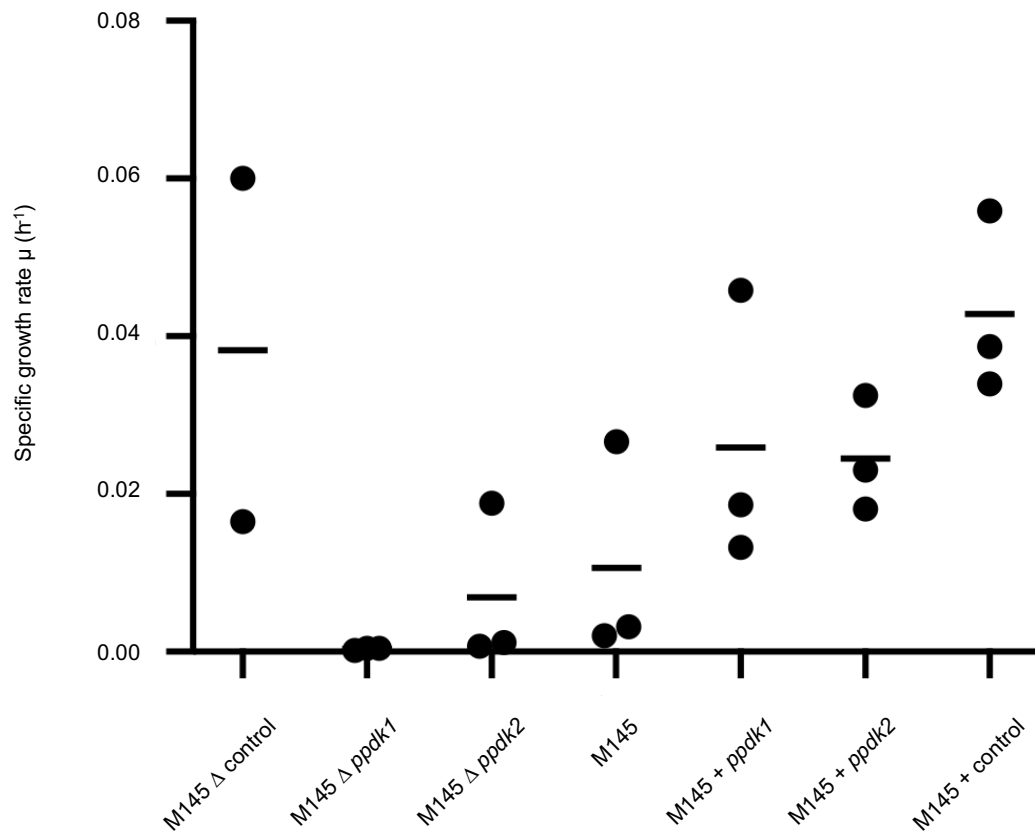


Figure 4.3 Specific growth rate of *S. coelicolor* M145 and mutants $\Delta/+$ (*ppdk1*; *SCO2028* and *ppdk2*; *SCO2494*) when grown on minimal media supplemented with glucose. Growth experiments were performed in triplicate and the specific growth rate calculated for each culture; these data are plotted as individual points on the graph with the mean growth rate illustrated by a horizontal line. Statistical testing was conducted using ANOVA however there were no statistically significant differences in growth rate between strains.

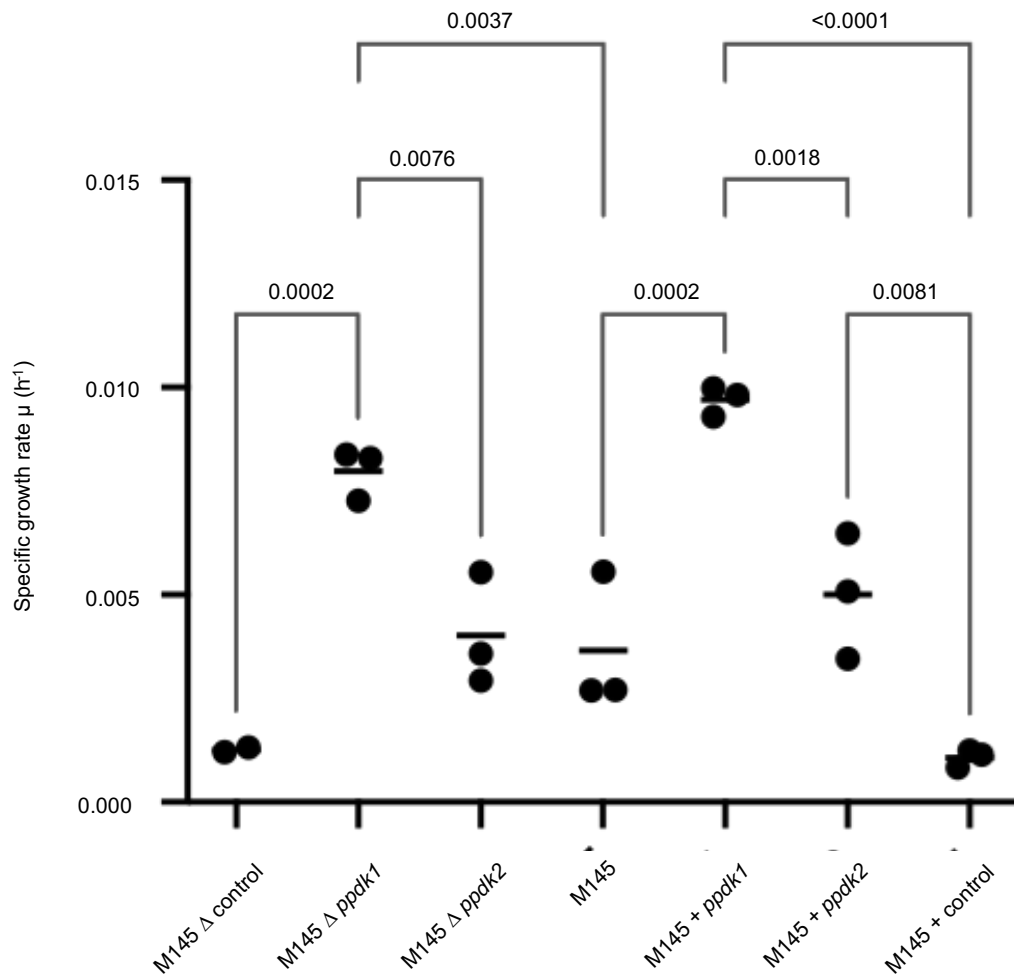


Figure 4.4 Specific growth rate of *S. coelicolor* M145 and mutants $\Delta/+$ (*ppdk1*; SCO2028 and *ppdk2*; SCO2494) when grown on minimal media supplemented with Tween 40. Each growth experiment was performed in triplicate and the specific growth rate calculated for each culture; these data are each plotted as an individual point on the graph with the mean growth rate of each strain illustrated by a horizontal line. Statistical analysis was performed in the form of ANOVA, and the p values stated.

When grown on minimal media supplemented with glucose (Figure 4.3), variations in specific growth rates are observed between the model *S. coelicolor* M145 and its mutant derivatives. Knocking down either *ppdk* (*ppdk1*; SCO0208, *ppdk2*; SCO2494) leads to a slight reduction in specific growth rate compared to the wild type. Conversely, overexpressing either *ppdk* (*ppdk1*; SCO0208, *ppdk2*; SCO2494) increases the specific growth rate ~2-fold. If *ppdk* catalyses the reaction of PYR-PEP in both directions it is possible that overexpression of *ppdk* is reducing the bottleneck of C₃ as it reaches the last stage of glycolysis and is increasing carbon flux in the glycolytic direction. Equally, knocking down *ppdk* reduced the growth rate suggesting that both *ppdk1* and *ppdk2* were involved in growth on glucose. This could mean that by knocking down either *ppdk* it is merely reducing the possible routes for carbon to be recycled in central metabolism. However ANOVA testing of the data did not provide any significant results to support this hypothesis.

When these strains were grown on minimal media supplemented with Tween 40 (Figure 4.4), there is an overall reduction in growth rate compared to that of glucose. This is unsurprising as it is a poor source of carbon when compared to glucose. The expression of *ppdk2* (SCO2494), either overexpression or knockdown, had no noticeable effect on growth rate when supplemented with Tween 40, suggesting it is not important for gluconeogenic growth. Conversely, the growth rate of both $\Delta ppdk1$ and *+ppdk1* increased significantly compared to the WT M145. This suggests that *ppdk1* is very much involved in controlling gluconeogenic growth, with the growth rate of the WT doubling when *ppdk1* is overexpressed.

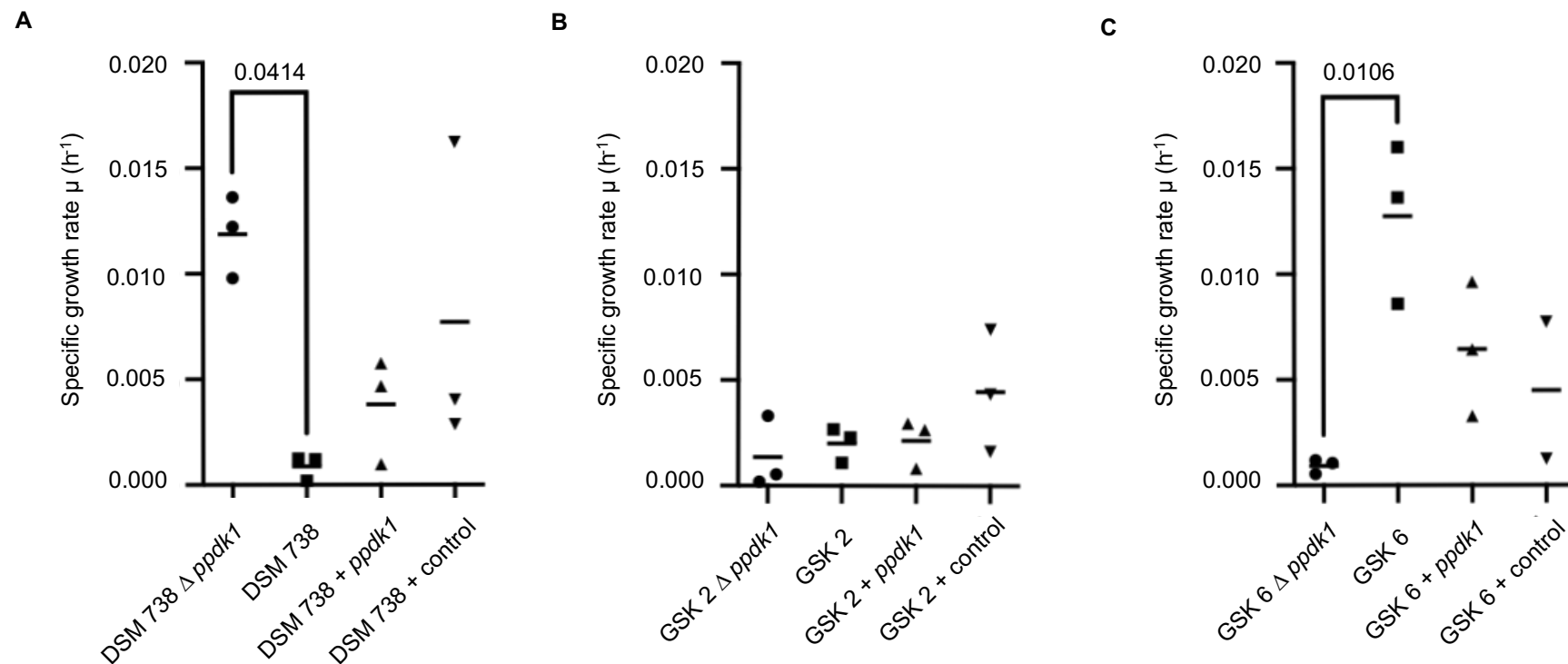


Figure 4.5 Growth of *S. clavuligerus* on minimal medium with glucose as the sole carbon source. A, B, and C show the specific growth rates of DSM738, GSK2, and GSK6 and their derivative strains respectively when grown on minimal media supplemented with glucose. Values were calculated from three independent cultures per strain, which are represented by a marker and the mean shown by a line. Statistical analysis was performed in the form of ANOVA.

When grown on minimal media supplemented with glucose the *S. clavuligerus* strains have lower specific growth rates compared to *S. coelicolor* strains. When glucose is the sole carbon source, *S. clavuligerus* DSM738 derivative strains have a higher specific growth rate than that of the parental strain. Figure 4.5 A shows a difference between DSM 738 and DSM 738 $\Delta ppdk$ (ANOVA) where $p=0.0414$. Knocking down *ppdk* (SCLAV1689) in DSM738 results in a >10-fold increase in specific growth rate (from $0.882 \times 10^{-3} \text{ h}^{-1}$ to $11.879 \times 10^{-3} \text{ h}^{-1}$). This suggests that knocking down *ppdk* in this instance is advantageous to growth.

For *S. clavuligerus* GSK2, the second industrial strain used in this thesis, a low growth rate ($\sim 0.002 \text{ h}^{-1}$) is observed in all strains with no significant differences in the growth rate between the parental and mutant strains.

The most recent industrial strain used in these experiments, *S. clavuligerus* GSK6, shows the highest specific growth rate of any of the *S. clavuligerus* strains with a specific growth rate of 0.013 h^{-1} . Figure 4.5 C shows a significant difference in growth rate between *S. clavuligerus* GSK6 and its derivative knockdown, *S. clavuligerus* GSK6 $\Delta ppdk$ ($p=0.0054$). The reduction in growth rate when *ppdk* is knocked down in *S. clavuligerus* GSK6 shows that when grown on glucose as a sole carbon source, *ppdk* is playing an important role.

If the three parental *S. clavuligerus* strains, DSM738, GSK2, and GSK6, are compared, there is an increase in growth rate on glucose between each subsequent strain although the difference between the Wild-Type *S. clavuligerus* DSM738 and *S. clavuligerus* GSK2 is only small there is a significant change in rate between the two industrial strains GSK2 and GSK6 ($p=0.0145$) implying that *S. clavuligerus* GSK6 is better adapted for growth on glucose than its predecessors.

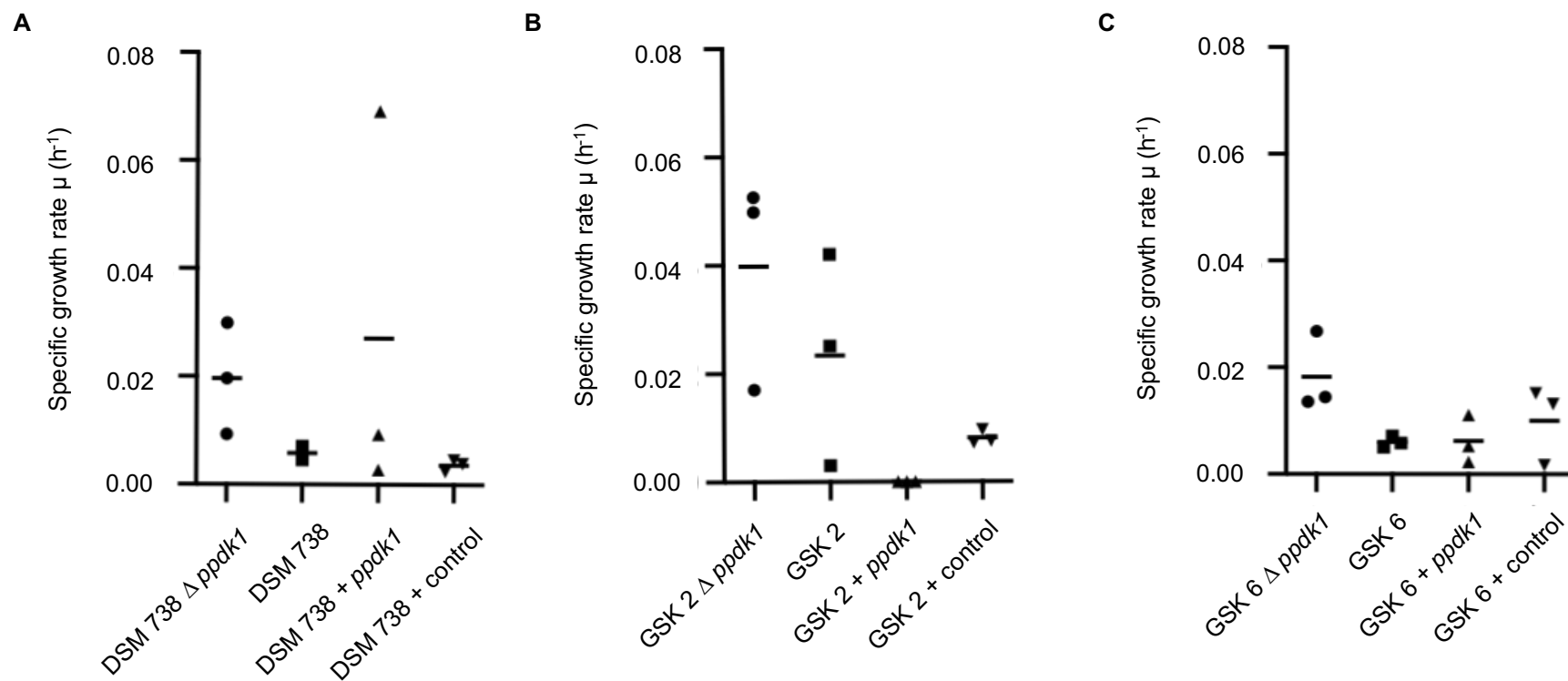


Figure 4.6 Growth of *S. clavuligerus* on minimal medium with Tween 40 as the sole carbon source. A, B and C show the specific growth rates of DSM738, GSK2, and GSK6 and their respective mutants respectively when grown on minimal media supplemented with Tween 40. Growth rates were calculated from three independent cultures per strain, which are each shown by a marker and the mean shown by a line. No statistical differences (ANOVA) were observed between these strains.

In contrast, when the *S. clavuligerus* strains are grown on minimal media supplemented with Tween 40 the growth rates are generally higher than those of *S. coelicolor* when grown on the same media. This likely reflects the fact that the industrial strains have been adapted over long periods of time to grow on similar oil-like carbon sources. In *S. clavuligerus* DSM738, a ~5-fold increase in growth rate is observed when *ppdk* is knocked down, similar to the response seen when the strains are grown on glucose. *S. clavuligerus* GSK2 has a higher mean growth rate than *S. clavuligerus* DSM738 but shows the same trend: when *ppdk* is knocked down in this strain the growth rate also increases. When *ppdk* is over expressed in this strain, it appears to almost abolish growth completely, suggesting that *ppdk* is limiting the growth rate of these strains when grown on oil-based media. Finally, a similar trend was seen in *S. clavuligerus* GSK6 and its mutants: when *ppdk* was knocked down the growth rate doubled but in this case over expressing *ppdk* in this strain had negligible effect on the growth rate. Overall, when the *S. clavuligerus* strains were grown on Tween 40 as a sole carbon source, they had higher specific growth rates than when grown on glucose. However, the changes in growth rate between the parental strains and the mutants when grown on Tween 40 were not as stark as those observed when grown on glucose. There was also a similar trend observed between the three parental strains and their mutants, which is that the growth rate increased when *ppdk* was knocked down.

4.5 Altering expression of *ppdk2* (SCO2494) in *S. coelicolor* impacts antibiotic titre

It is clear that altering expression of *ppdk* can impact growth rate, which varies depending on the strain and the primary carbon source. However, growth rate is not a linear proxy for antibiotic production, which is an essential operational parameter for the industrial strains. Thus, it was essential to specifically investigate the role of *ppdk* in antibiotic production under the same growth conditions. To determine this, the antibiotic titres of the *Streptomyces spp* and their mutant strains were measured after being grown on minimal media supplemented with either glucose or Tween 40 until the culture reached stationary phase.

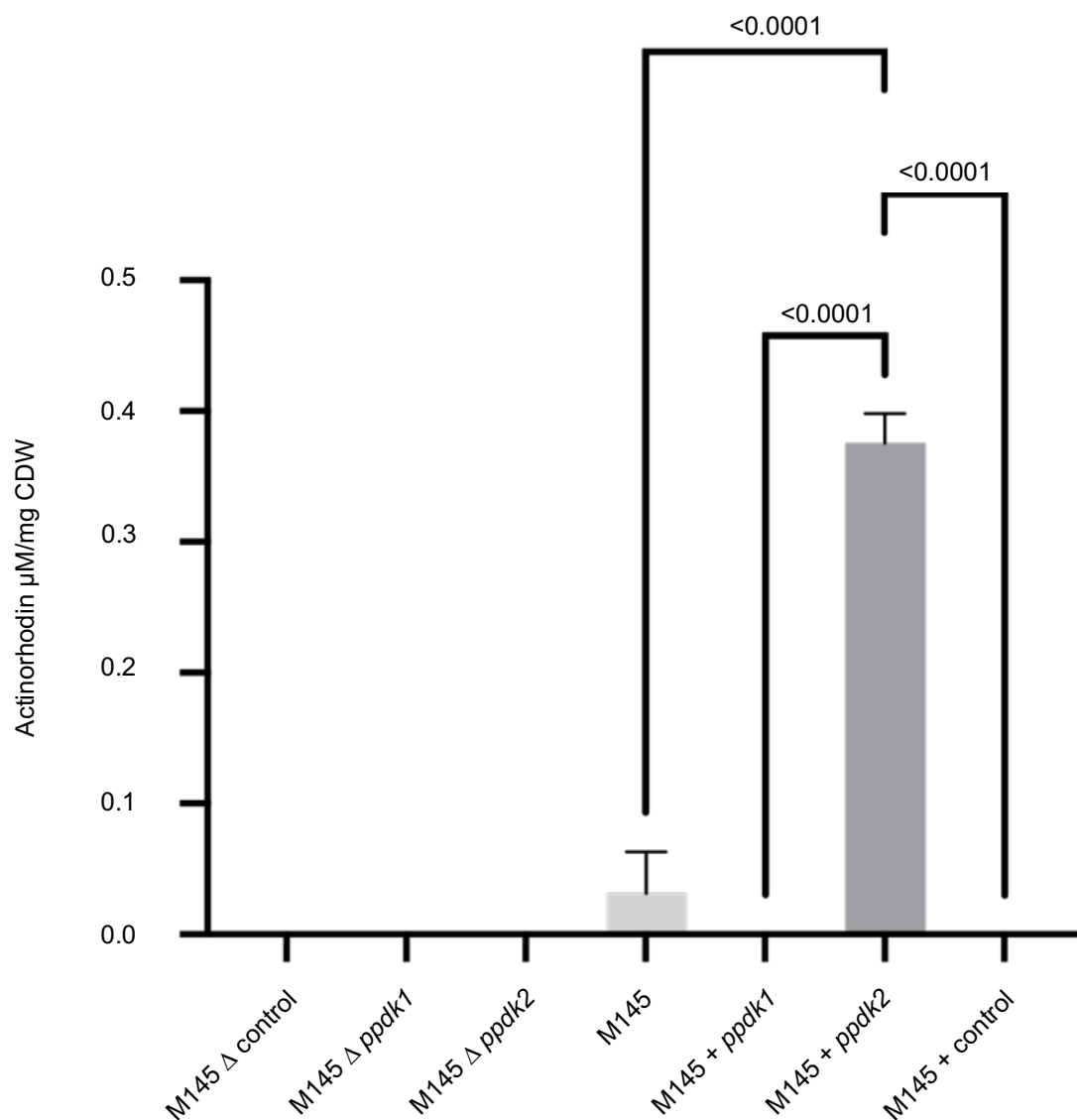


Figure 4.7 Actinorhodin titre of *S. coelicolor* and isogenic mutant strains when grown on glucose. Average titre (of three replicates) plotted, and error bars show standard error of the mean. The graph shows knocking down either *ppdk* gene abolishes antibiotic production when grown on this media. An increase (ANOVA) in actinorhodin production is observed in the M145 +*ppdk2* mutant compared to the Wild-Type M145 strain ($p < 0.0001$).

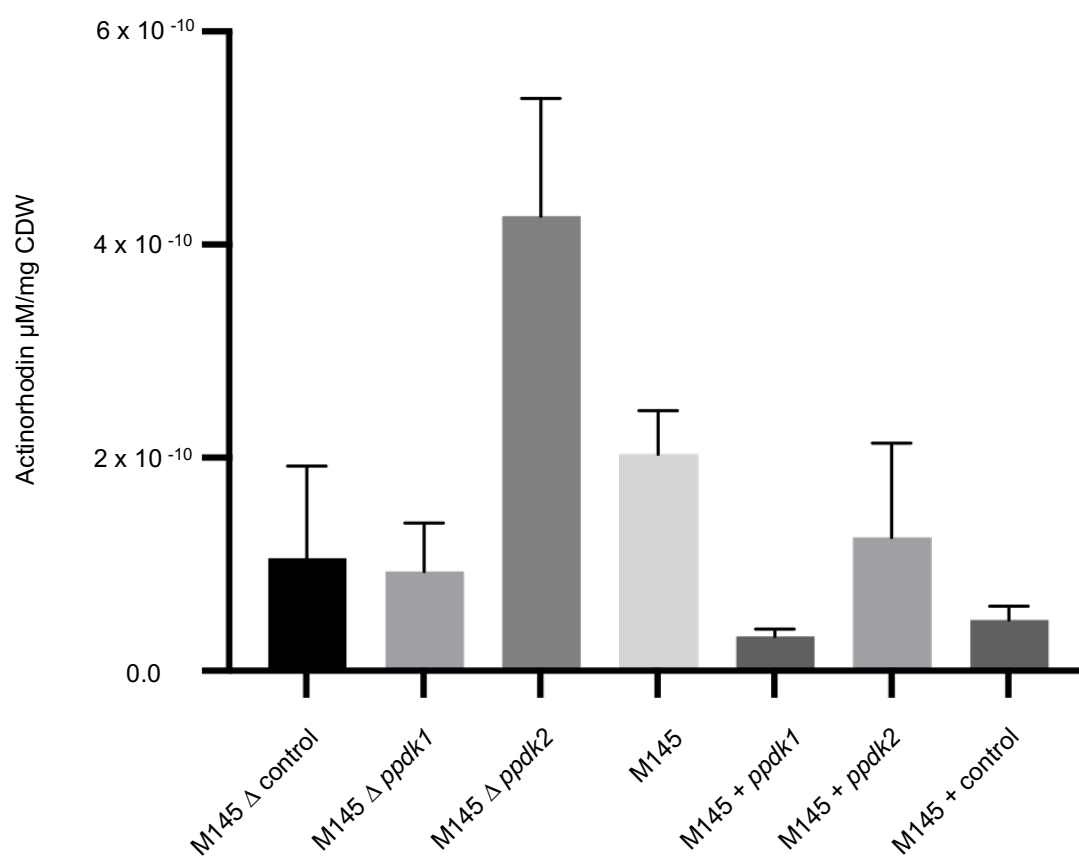


Figure 4.8 Actinorhodin titre of *S. coelicolor* and derivative mutants strains when grown on Tween 40 as a sole carbon source. Average titre (of three replicates) plotted and error bars show standard error of the mean.

When the strains were grown on glucose it was observed that knocking down either *S. coelicolor ppdk1* or *ppdk2* (SCO0208 and SCO2494) completely abolished actinorhodin production (Figure 4.7). This is unsurprising as both knock down strains had a decreased growth rate compared to the Wild-Type strain when grown on glucose as a sole carbon source and when these genes were overexpressed an increase in growth rate was observed.

When the *ppdk2* overexpression mutant is grown on glucose, the antibiotic titre is significantly increased compared to the Wild-Type, however this is not true for *ppdk1*. This shows that the increase in antibiotic titre observed is not just an artefact of the increased growth rate and that *ppdk2* is important for the production of actinorhodin when grown on glucose.

In contrast to minimal media supplemented with glucose when *S. coelicolor* M145 and the isogenic mutant strains were grown on Tween 40 as a sole carbon source the opposite is observed (Figure 4.8). This time when knocking down *ppdk2*, an increase in actinorhodin production was observed compared to that of the *S. coelicolor* Wild-Type. Over expressing *ppdk1* on the glucose media appeared to decrease antibiotic titre compared to the *S. coelicolor* Wild-Type but other strains had unaltered actinorhodin titres. If the titres of actinorhodin and growth rates of these *S. coelicolor* strains on Tween 40 are compared, it shows that the strains producing a higher titre of the antibiotic, *S. coelicolor* $\Delta ppdk2$, are not the strains growing at an increased rate, *S. coelicolor* $\Delta ppdk1$ and $+ppdk1$. This suggests that the resulting increase in titre is independent of the rate of growth and that in the case of actinorhodin production *S. coelicolor ppdk2* (SCO2494) appears to play a role in antibiotic production.

Low titres of CA were expected due to the nature of the sole carbon source growth conditions but combined with the low sensitivity of the method used for the

measurements it was unfortunately not possible to measure and therefore report the levels CA produced by any of the *S. clavuligerus* strains.

4.6 Interspecies complementation experiments

Complementation experiments carried out by Professor Paul Hoskisson investigate the interconversion of PEP and pyruvate which has been well studied in *E. coli* but neglected in *Streptomyces* [140]. *E. coli* lacks PPDK and to grow gluconeogenically on pyruvate requires the activity of phosphoenolpyruvate synthase (PPS). PPS converts pyruvate to phosphoenolpyruvate (PEP) when pyruvate is the sole carbon source. *E. coli* possesses two copies of pyruvate kinase (*pykA* and *pykF*), that convert PEP to pyruvate. In *Streptomyces* both of these reactions can be carried out by PPDK. PykF is the primary enzyme responsible for this in *E. coli*, with PykA playing a minor role. Schniete *et al.*, demonstrated that $\Delta pykA$ and $\Delta pykF$ mutants and a $\Delta pykAF$ double mutant could be complemented by the paralogous pyruvate kinases from *Streptomyces* [40].

To investigate the physiological role of the *ppdk* genes from *S. coelicolor* and *S. clavuligerus* and to establish if they are able to function gluconeogenically and/or glycolytically given they can reversibly catalyse the pyruvate to PEP reaction a cross-species complementation experiment was undertaken.

The Δpps mutant from the Keio collection [140] was obtained from Dr Manuel Banzhaf (University of Birmingham) from an ordered and verified Keio Collection library. The $\Delta pykAF$ double mutant was obtained from the strain collection in the Hoskisson Laboratory. To each of these strains the paralogous *ppdk* genes (encoded by *SCO0208* and *SCO2494*) from *S. coelicolor* and the orthologous *ppdk* from *S. clavuligerus* (*SCLAV1689*) were introduced. The strains were grown overnight from a

single colony, along with the isogenic parent strain (BW25113) in 5 ml of LB broth. The overnight culture was diluted 1/100 and grown in LB to an OD₆₀₀ of 0.4, before washing in M9 salts to remove residual carbon sources from growth on complete M9, or direct transfer in the case of LB. Growth curves were initiated with 1% (v/v) of exponentially growing culture.

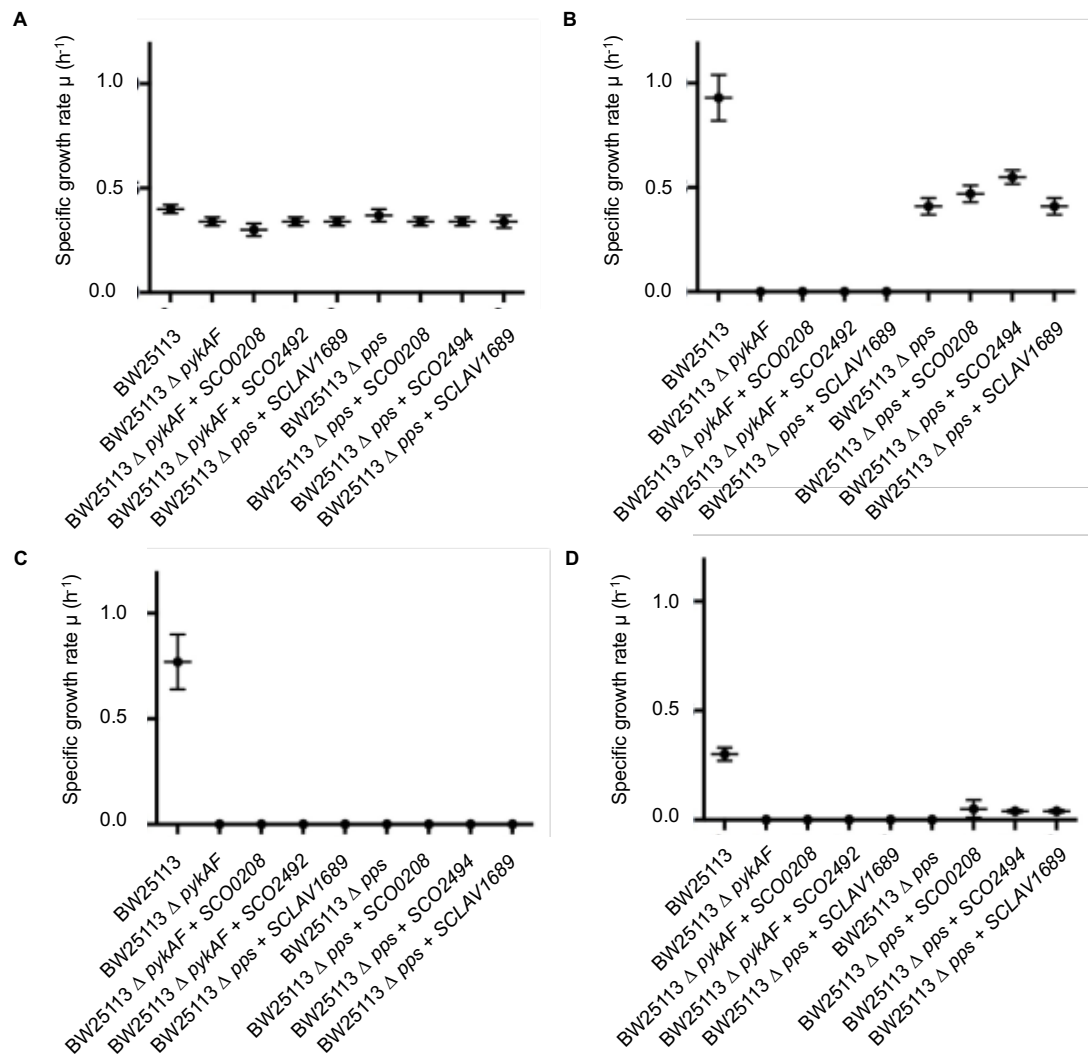


Figure 4.9 Cross species complementation of the PEP to pyruvate metabolic node. Specific growth rates of *E. coli* isogenic parent strain (BW25113), the Keio Collection *pps* mutant and the pyruvate kinase double mutant [40]. Strains were grown in LB (A), M9 medium plus either glucose (B; 1%), pyruvate (C; 0.4%) or PEP (D; 0.4%).

To establish the physiological role of the PPDK enzymes in an *E. coli* genetic background, all strains were grown on LB as a control for growth on complex medium, and all strains were grown on LB and M9 medium plus a carbon source to drive a particular a node of carbon metabolism. To drive cells to grow glycolytically, glucose (1%) was used as a sole carbon source. To investigate the ability of the *Streptomyces* PPDK enzymes to support gluconeogenic growth (the presumed role of PPDK), pyruvate (0.4%) was used as a sole carbon source and finally to investigate if the PPDKs can catalyse the reverse reaction (glycolytic) under physiological conditions phosphoenolpyruvate (0.4%) was used as a sole carbon source.

When all strains (Parent [BW25113], $\Delta pykAF$, $\Delta pykAF + SCO0208$, $\Delta pykAF + SCO2494$, $\Delta pykAF + SCLAV1689$, Δpps , $\Delta pps + SCO0208$, $\Delta pps + SCO2494$ and $\Delta pps + SCLAV1689$) were grown in LB medium (Figure 4.9 A), no difference was observed between the growth of the strains. This is likely due to the potential to grow using a variety of carbon sources under these conditions, with strains not relying on a single metabolic pathway to grow.

When all strains were grown on M9 medium with glucose (1%) as the sole carbon source, the $\Delta pykAF$ mutant and all of the derivatives ($\Delta pykAF + SCO0208$, $\Delta pykAF + SCO2494$, $\Delta pykAF + SCLAV1689$) were unable to grow (Figure 4.9 B). This is consistent with the report of Schniete *et al.*, [40] that a $\Delta pykAF$ mutant was unable to grow when glucose was a sole carbon source. Moreover, these data indicate that the *ppdk* genes from *Streptomyces* are unable to complement this genetic lesion, and that under these physiological conditions PPDK does not function in the reverse direction (PEP to pyruvate). The Δpps strain and its derivatives ($\Delta pps + SCO0208$, $\Delta pps + SCO2494$ and $\Delta pps + SCLAV1689$) were all able to grow using glucose as a sole carbon source, although at a rate of about half that of the isogenic parent strain BW25113. Under these conditions there is unlikely to be much gluconeogenic flux.

To investigate the role of the *Streptomyces ppdk* genes when pyruvate was the sole carbon source, all strains were grown on M9 medium with pyruvate (0.4%; Figure 4.9 C). Under these conditions, none of the strains other than the parental strain (BW25113) were able to grow. This suggests that the gluconeogenic conversion of pyruvate to PEP by PPDK is not possible in these strains or under these physiological conditions. In *E. coli*, *pps* null mutants are known to be unable to grow on pyruvate as a sole carbon source [140]. These data suggest that the $\Delta pykAF$ mutant also does not grow well when pyruvate is the sole carbon source, likely due to insufficient carbon for anaplerosis.

To further investigate the *ppdk* genes, all strains were grown on M9 medium with phosphoenolpyruvate (0.4%) as the sole carbon source. It was hypothesised under these conditions that the $\Delta pykAF$ strain and all derivatives would be unable to grow, mimicking the phenotype observed on glucose, and this was found to be the case (Figure 4.9 D). In the case of the Δpps strain, it was hypothesised that this mutant and its derivative would also behave like the glucose grown cultures, with all strains exhibiting growth. This was found not to be the case. The *pps* null mutant was unable to grow and a small amount of growth was observed in the complemented strains. This was unexpected as when grown with glucose, all strains were able to grow. Interestingly, during growth, normally PPS competes with pyruvate dehydrogenase (PDH) for pyruvate. PDH is responsible for the production of formation of acetyl-CoA, mediating entry of carbon into the citric acid cycle [140]. It is possible that a potentially small amount of gluconeogenic activity from the PPDKs under these conditions is able to facilitate the limited growth of the cells under these conditions.

Overall, it seems that the *Streptomyces ppdk* genes are unable to complement the *pps* phenotype in *E. coli*, carrying out the conversion of pyruvate to PEP under these conditions. It is possible that there could be missing effectors/cofactors required by

the enzymes combined with possibly unsuitable physiological conditions for *Streptomyces* enzymes to be expressed and/or to function in *E. coli*. The reverse reaction (pyruvate to PEP) while known to occur *in vitro* also was not possible in *E. coli* with the *Streptomyces* enzymes. Due to the lack of information about PPDK enzymes in general it would be remis to conclude that *ppdks* from *Streptomyces* cannot complement *pps* in *E. coli* without first screening other physiological conditions, particularly temperature.

4.7 Discussion

Currently around two-thirds of our clinically used antimicrobial drugs are produced by the bacterial genus *Streptomyces*. These are currently made industrially via fermentation using strains that have been through a decades-long ordeal of successive rounds of mutation to select for ever-improved titres and more economical production. Traditionally, these strain improvement efforts have relied on random mutagenesis approaches using chemical mutagens, followed by screening for an increase in antibiotic titre. This process can generate a wide range of enhancing mutations in industrial *Streptomyces* strains including: increase in biosynthetic gene expression, the loss of competing specialised metabolite expression, and increased precursor supply via primary metabolism or through increased adaptation to fermentation and media conditions. Whilst these can be advantageous, they can be detrimental later down the line if, for example, it is required that the strains are grown on a more sustainable feedstock and they have lost the ability to utilise varied carbon sources. Linking genotype to phenotype in these industrial strains remains a challenge as many of the effects of the random mutagenesis remain unknown. This Chapter aimed to try and lay out how the industrial strains differ phenotypically from the model and how the successive industrial strains differ from each other. To also

understand if the specific mutations made in Chapter 2 have impacted the phenotype of the mutant strains.

4.7.1 Phenotypic analysis reveals strains are adapted to industrial media.

Phenotypic analysis of WT *S. clavuligerus*, the industrial lineage strains and the derivative strains produced in this work show different growth patterns dependent on the medium. The parental strain of *S. clavuligerus* DSM738, and the industrial strains of *S. clavuligerus* GSK2, and GSK6 and their corresponding derivatives when grown on industrial media show an increasing adaptation to growing on these complex industrial media, which is unsurprising given the industrial development of these production strains. The process of strain improvement generally results in a range of responses in industrial strains from loss of competing specialised metabolite expression, increased production of natural products or increased stress resistance/sensitivity [40, 141, 142]. Improvement to industrial strains may occur at the level of biosynthesis, at the level of metabolism or through increased adaptation to fermentation and media conditions [143]. When an organism adapts genetically to one environment, they often lose fitness in other environments – essentially representing a form of ecological specialisation [137].

Schniete *et al.*, demonstrated that PPK played an important role in metabolism during the growth of strains on media that mimicked the oil-based media of industry [144]. CRISPR dCas9 disruption of *ppk* in strains *S. clavuligerus* DSM738 and GSK6 resulted in mutants that show limited growth, compared to the parental strains, but in *S. clavuligerus* GSK2 it abolishes growth. When *ppk* is knocked down this results in carbon flux being disrupted in the PEP-Pyruvate-oxaloacetate node of central metabolism [107] which limits gluconeogenic reactions that are known to be important during antibiotic biosynthesis [40].

When *ppdk* is overexpressed in *S. clavuligerus* GSK2, unlike in *S. clavuligerus* DSM738 and *S. clavuligerus* GSK6, growth is drastically increased suggesting that gluconeogenic carbon flux is important for biomass generation and may also increase the pool of precursor molecules for the biosynthesis of CA.

The stark contrast between the growth of *S. clavuligerus* GSK2-derived strains and those derived from *S. clavuligerus* DSM738 and GSK6 on the industrial media suggests that the accumulation of mutations in these strains not only hones metabolism for increased production but also results in deleterious mutations, reducing the capacity to catabolise a wide range of substrates as strains adapt to a particular medium.

4.7.2 Growth analysis of strains

To look at adaptation, the change in catabolic function during the strain improvement process was investigated, along with the effect of disrupting the *ppdk* genes (*SCO00208*, *SCO2494*, and *SCLAV1689*). It was hypothesised that unused catabolic function will decay due to mutation and substrate utilisation will become narrower. Using the OmniLog Biolog, an overview of how the different strains utilise different carbon sources was obtained. *S. coelicolor* M145 unsurprisingly utilised the widest range of carbon sources as this is the model organism and has not been adapted to an industrial setting. Of the *S. clavuligerus* strains, DSM738 unsurprisingly utilises the widest range of carbon sources compared with the industrial strains. However, altering the expression of *ppdk* in this strain still decreases the substrate utilisation. This diversity of available catabolic pathways further decreases as the strains move further along the industrial lineage until the main sources left being utilised by industrial strain GSK6 are the lipid like polysorbates Tween 40 and Tween 80 and N-acetyl glucosamine, the monomer of chitin and carbon source that would be utilised

in the natural environment as it is abundant [145]. Adaption of strains to a specific growth medium is a key component of industrial *Streptomyces* strain development, where loss of traits such as carbon catabolite repression of antibiotic biosynthesis and ability to utilise cheap carbon sources such as oil are beneficial to industrial production and process scale-up. Understanding how strains adapt to their environment will enable rapid future strain design for industrial strains of *Streptomyces* through recapitulation of key adaptive mutations in strains that are new to industrial optimisation. Whilst this has not been studied in industrial strains of *Streptomyces*, there is precedent for loss of substrate utilisation in adapted populations of *E. coli* [137].

4.7.3 Physiological role of *ppdk*

When growth and antibiotic production are investigated more closely using minimal media supplemented with either glucose or Tween 40 we see in *S. coelicolor* $\Delta ppdk1$ an increase in growth rate on Tween 40 which is also observed in *S. clavuligerus* $\Delta ppdk$. Although *S. coelicolor* *ppdk2* is the orthologous gene to the *S. clavuligerus* *ppdk*, we do not observe these similarities when the strains are grown on either carbon source. Instead, it appears that the *S. coelicolor* protein PPDK2 is regulating antibiotic production as when it is overexpressed during growth on glucose, we see an increase in antibiotic production, and the opposite is observed when grown on Tween, when it is knocked down we see an increase in antibiotic production. It is possible that when *ppdk2* is knocked down in *S. coelicolor* and the strain is grown on glucose that the increase in actinorhodin production observed is because a gluconeogenic pathway for growth has been removed. This could mean that rather than carbon flux being pushed in the gluconeogenic direction it was instead being processed through the TCA cycle and the Acetyl-CoA being syphoned off for the production of antibiotics.

This is not the first time something like this has been done: phosphofructokinase, *pfkA*, and the genes *zwf1* and *zwf2* encoding the glucose-6-phosphate dehydrogenase enzymes have previously been targeted to reduce carbon flux through the pentose phosphate pathway, another gluconeogenic pathway, to also increase precursor availability of Acetyl-CoA for actinorhodin production [144].

It was not possible to draw any conclusions about the antibiotic titre from the *S. clavuligerus* strains as the readings were below the threshold of our assay, so for future investigation a more sensitive method would be required, for example using HPLC to evaluate concentrations rather than a plate reader assay. There has also been research on flux around this anaplerotic node suggesting that carbon limited conditions can reduce flux and therefore precursor availability for CA production [101]. These findings suggest that PPDK enzymes do affect flux around the anaplerotic node, which impacts both growth and antibiotic production.

Engineering *Streptomyces* with mutations that facilitate increased fitness and antibiotic production in different genetic backgrounds has revealed a lot about how industrial strains have adapted to production processes. These data provide a starting point to develop in depth knowledge of epistasis across the genomes of industrial strains and framework that can inform on the generation of industrial strains with better production characteristics. However, what can be concluded from the experiments in this Chapter is that they suggest the two PPDK enzymes from *S. coelicolor* do not function in the same way and this could be important for controlling carbon flux. To try and further the understanding of these primary metabolic enzymes, they will be investigated in Chapter 5.

**Chapter 5: Expression, purification,
and characterisation of Pyruvate
phosphate dikinase proteins from
S. clavuligerus and *S. coelicolor***

5.1 Introduction

PPDK enzymes are found across all domains of life and play a role in primary metabolism. In plants, PPDK enzymes play a major role in the Calvin cycle, generating PEP to act as the major CO₂ acceptor [146]. In some bacteria and archaea where pyruvate kinase is absent, PPDK enzymes bridge the gap between glycolysis and the TCA cycle enabling organisms to grow gluconeogenically through the conversion of pyruvate to PEP [147]. In some major families of organisms, such as the Enterobacteriales, this activity is replaced by the irreversible enzyme phosphoenolpyruvate synthase (PPS) [140]. In organisms where pyruvate kinases are present, it is generally considered that the enzymes are more active in the forward PEP-producing direction. In *Streptomyces*, where there are two pyruvate kinases, it is likely that these PPDK enzymes function in the gluconeogenic direction (pyruvate to PEP) [40]. There is further evidence to support this, where it was shown that when grown on gluconeogenic substrates, *ppdk* in *S. coelicolor* was upregulated around 30-fold [40, 99]. In Chapter 4 it was shown that CRISPRi knock downs of the *ppdk* genes in two *Streptomyces* species affected growth and antibiotic production under several different conditions.

Although these genes are found in numerous organisms, few PPDK enzymes have been characterised and most of these studies have been in plants, more specifically maize and other C₄ organisms [148]. As discussed in the introduction, *ppdk* genes are found abundantly in Actinobacterial genomes, unlike in other bacterial species, yet there have been limited functional studies involving these organisms. One of the few available studies revealed that, in the thermophilic Actinomycete *Microbispora rosea*, PPDK is ~90 kDa, and reported apparent *K_m* values (5 µM, 38 µM and 280 µM for AMP, PP_i, and PEP respectively) which are comparable to those of other organisms [147]. There have also been a number of studies involving the Gram-positive

bacterium *Clostridium symbiosum* which have reported K_m values in the range of 15 μM , 80 μM , and 100 μM for AMP, PP_i , and PEP respectively [149, 150]. Although these two studies are notable, for the most part, PPDK research has not focussed on bacterial enzymes which is therefore a crucial knowledge gap in the bacterial physiology of organisms that have these genes. The direction of function of PPDK can determine increased flux of carbon used to produce specialised metabolites such as actinorhodin as shown in Chapter 4. The phylogenetic tree in the Introduction (Figure 1.7) illustrates just how many *ppdk* genes are present in actinobacterial species. These genes provide many potential targets for genetic manipulation which could be used to increase the production of high value specialised metabolites. The aim of this Chapter was to establish a system for heterologous expression and subsequent purification of active PPDK proteins from *S. coelicolor* and *S. clavuligerus*. Once purified, the proteins would then be characterised to provide insight into their role within *Streptomyces* metabolism and investigate functional differentiation of PPDK between the strains.

5.2 Specific Objectives:

- To predict the structures of the PPDK proteins from *S. clavuligerus* and *S. coelicolor* using AlphaFold2.
- To heterologously express the PPDK protein from *S. clavuligerus*, an industrially relevant strain and the two PPDK enzymes from *S. coelicolor*, the model species of *Streptomyces*.
- To optimise the purification of his-tagged PPDK proteins using Immobilised Metal Affinity Chromatography (IMAC).
- To characterise PPDK proteins by a preliminary study of their enzyme activities.

5.3 PPDK sequence is conserved throughout the industrial lineage of *S. clavuligerus*.

The amino acid sequences of the three *S. clavuligerus* PPDK proteins (encoded by *SCLAV1689*) from strains DSM738, GSK2 and GSK6 were obtained from GSK and aligned using Clustal Omega [151] to establish their similarity. The resulting alignment (Figure 5.1) shows that the sequences are completely conserved throughout the lineage. The two industrial strains show 100 % identity when aligned to the Wild-type strain DSM738 used as the reference strain [152]. The fact that they are so highly conserved despite many rounds of mutagenesis suggests that retaining the PPDK enzyme confers some advantage to the strains.

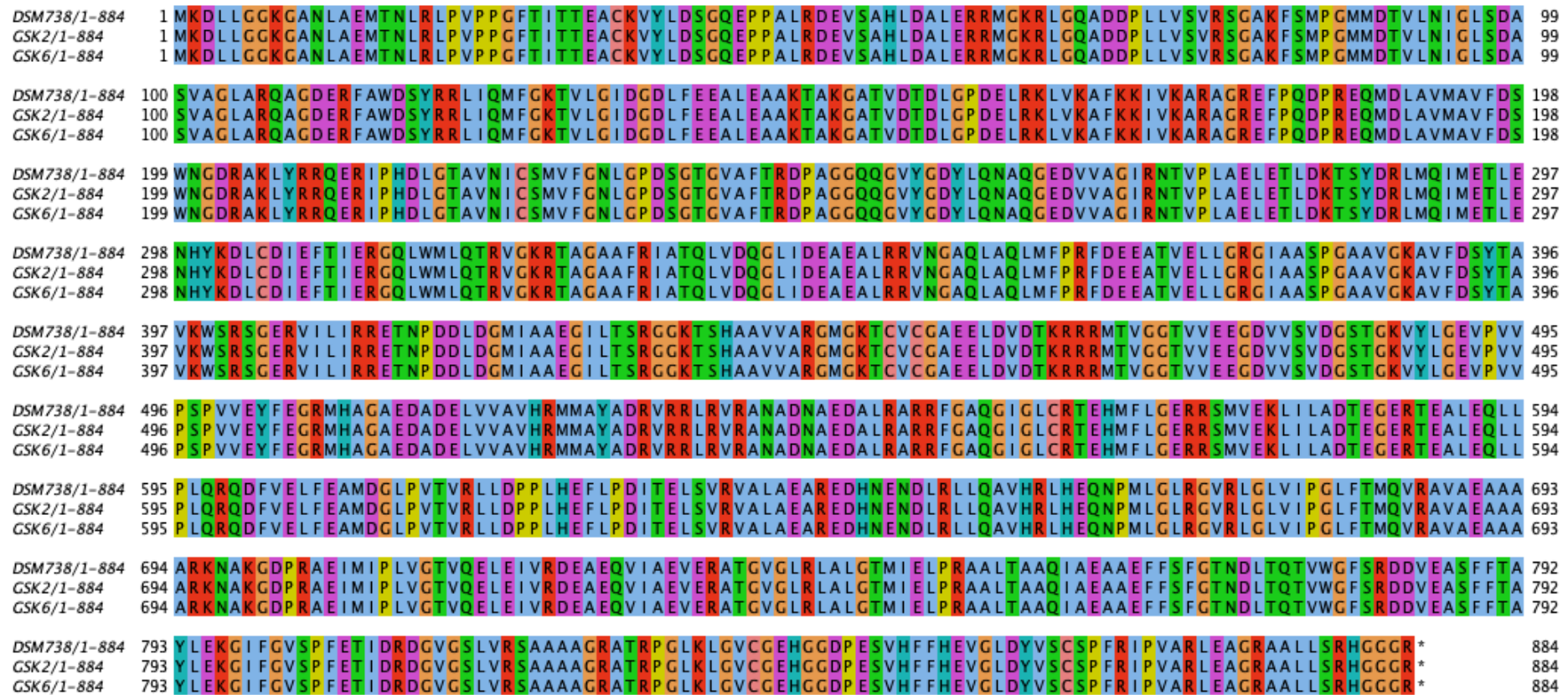


Figure 5.1 Multiple sequence alignment of PDK (SCLAV1689) amino acid sequences from *S. clavuligerus* strains. Amino acid sequences aligned using Clustal Omega come from the Wild-Type strain DSM738 and the two industrial strains GSK2 and GSK6.

5.4 Comparison of *Streptomyces* PPDK sequences and sequences of previously characterised PPDK proteins.

PPDK proteins are found in a range of organisms including plants, archaea, and bacteria, but the crystal structures of only three PPDK proteins have been resolved: these are from *Trypanosoma brucei*, *Zea mays*, and *Clostridium symbiosum* [153–155]. To investigate conservation of these proteins, the amino acid sequences of the PPDK proteins from *S. coelicolor*, *S. clavuligerus*, *T. brucei*, *Z. mays*, and *C. symbiosum* were aligned using Clustal Omega.

When aligned to the Wild-Type *S. clavuligerus* strain DSM738 the other *Streptomyces* PPDK amino acid sequences show 100% identity and the *Z. mays*, *C. symbiosum*, and *T. brucei* sequences show 96.3%, 96.6%, and 96.8% similarity respectively. The variation in sequence similarity is to be expected as the organisms from which these proteins originate are not closely related. The binding sites and two active sites of the PPDK proteins from *T. brucei* [156], *Z. mays* [157], and *C. symbiosum* [149] are documented in the literature. These sites are highlighted in blue (binding site) and red (active site) in Figure 5.2. The sites are completely conserved throughout all species.

SCLAV_5441/1-609 483 LDHLGLAAEGVVTARGGSGHAAVVARSMGKPAVVGAAGLTVDVAAACVRA--GGHIVPEGTVIALDGTGGEVVVGEPVVTPAADA----- 568
 DSM738/1-905 438 DDLDGMIAAEGILTSRGGKTSAAVVARGMGKTCVCGAEELVDTKRRRMTV--GGTVVEEGDVSVSDGSTGKVYLGVEVPVPSPVVEYFEGRMHAGAEADEL 539
 GSK2/1-905 438 DDLDGMIAAEGILTSRGGKTSAAVVARGMGKTCVCGAEELVDTKRRRMTV--GGTVVEEGDVSVSDGSTGKVYLGVEVPVPSPVVEYFEGRMHAGAEADEL 539
 GSK6/1-905 438 DDLDGMIAAEGILTSRGGKTSAAVVARGMGKTCVCGAEELVDTKRRRMTV--GGTVVEEGDVSVSDGSTGKVYLGVEVPVPSPVVEYFEGRMHAGAEADEL 539
 SCO0208/1-898 430 DDLPGMIAAQAVLT SRGGKTSAAVVARGMGKTCVCGAEELVDTKARRCTAPDGTIVEEGTVLSVDGTAGTVYAGAVPLVDSAVMRYLEAGA--GGRATTGV 530
 SCO2494/1-909 441 DDLDGMIAAEGILTSRGGKTSAAVVARGMGKTCVCGAEELVDTKRRRMTVPGGHVVEEGDVISIDGSSGKVYLGVEVPVPSPVVEYFEGRMHAGAEADEL 543
 Zea/1-876 437 EDVGGMHAAVGILT ERGGMTSAAVVARGMGKCCVSCGSGIRVNDAEKLVTI--GGHVLREGEWLSLNGSTGEVILGKQPLSPALSG----- 522
 Trypanosoma/1-913 462 EDLAGMDAACGILTARGGMTSAAVVARGMGKCCVSCGCGDMVIR--GKSFKL--NGSVFREGDYITIDGSKGLIYAGKLLRSPDLKG----- 545
 Clostridium/1-874 434 EDIEGMHAAEGILTVRGGMTSAAVVARGMGTCVSCGCGEIKINEEAKTFEL--GGHTFAEGDYISLDGSTGKIYKGDIEQEAESVSG----- 519

 SCLAV_5441/1-609 569 --RLHRLLLAWADGFAGR--PPGGPADGTWAGRLAAQAVLRIG*----- 609
 DSM738/1-905 540 VVAVHRMMAYADRVRLRVANADNAEDALRRRFGAOGIGLCRTEHMFELGE--RRSMVEKLIADTEGERTEALEQLLP LQRQDFVELFEAMDGLPVTVRL 640
 GSK2/1-905 540 VVAVHRMMAYADRVRLRVANADNAEDALRRRFGAOGIGLCRTEHMFELGE--RRSMVEKLIADTEGERTEALEQLLP LQRQDFVELFEAMDGLPVTVRL 640
 GSK6/1-905 540 VVAVHRMMAYADRVRLRVANADNAEDALRRRFGAOGIGLCRTEHMFELGE--RRSMVEKLIADTEGERTEALEQLLP LQRQDFVELFEAMDGLPVTVRL 640
 SCO0208/1-898 531 VDAAVARALARADAVRRLRVANADTPQDAVRRRFGAOGIGLCRTEHMFELGE--RRKLVEMILAGTDAGRERALLALLP LQREDFTGILEAMDGLPVTI 631
 SCO2494/1-909 544 VEAVHRIMAFADRRLRVANADNAEDALRRRFGAOGIGLCRTEHMFELGD--RRLEVERLIADTEAVRREESLKEELLPLQKQDFVQLFESMDGLPVTVRL 644
 Zea/1-876 523 --DLCTFMAWVDYRKLVLANADTPDDALTARNNGAOGIGLCRTEHMFASDERIKAVRQIMAPTLERQQALDRLLPYQRSDFEGLFRAMDGLPVTI 623
 Trypanosoma/1-913 546 --SFTQILQWCQEMKRLGVRTNADTPADAAKARSFGAEGVGLCRTEHMFEGS--RINFIREMILADSASGRKAALDKLLP IQRADFVGLLRAMRCLPVTI 645
 Clostridium/1-874 520 --SFERIMVWADKFRTLKVRTNADTPEDTLNAVKLGAEGIGLCRTEHMFEEAD--RIMKIRKMLSDSVEAREEALNELIPFQKGDFAKMYKALEGRPMTV 619

 SCLAV_5441/1-609 641 DPLLHEFLPDITELSVRVALAEARE--DHNENDLRLQAVHRLHEQNPMLGLRGVRLGLVLPGLFTMQVRAVAEAAAARKN--AKGDPRAEIMIPLVGTVQLEI 741
 DSM738/1-905 641 DPLLHEFLPDITELSVRVALAEARE--DHNENDLRLQAVHRLHEQNPMLGLRGVRLGLVLPGLFTMQVRAVAEAAAARKN--AKGDPRAEIMIPLVGTVQLEI 741
 GSK2/1-905 641 DPLLHEFLPDITELSVRVALAEARE--DHNENDLRLQAVHRLHEQNPMLGLRGVRLGLVLPGLFTMQVRAVAEAAAARKN--AKGDPRAEIMIPLVGTVQLEI 741
 GSK6/1-905 632 DPLLHEFLPDRLADLAVRIATAEARGEEDDAHDAELLAAVGRMHENPMLGLRGVRLGLVLPGLVAMQVRALEGAQAQLT--AGGAPRAEIMVPLVGDVRELRL 733
 SCO0208/1-898 645 DPLLHEFLPDITELSVRVALAESRQ--EPHENELRLQAVHRLHEQNPMLGLRGVRLGLVLPGLFTMQVRAIEAAAERKA--AKGDPRAEIMIPLVGTVQLEI 745
 SCO2494/1-909 624 DPLLHEFLPEGNIEDIIVSELCAET---GANQEDALARIEKLSVNPMLGFRGCRGLGISYPELTQMARAIEAAIAMTN--QGQVFPFIMVPLVGTPELQELGH 721
 Zea/1-876 646 DPLLHEFVPHDAAAQ--FELAQKL---GMPAEKVRNRVNALHELNPMLGHRGCRGLITYPEIYNMQVRAIEAAIAVSE--EGSSVIFIMVPLVGKKEELSL 741
 Trypanosoma/1-913 620 DPLLHEFVPHTEEEQ--ALEAKNM---GLTLAEVKKAVDELHEFNPMGHRGCRGLAVTYPEIAKMQTRAVMEAAIEVKEETGIDIVPEIMIPLVGEKKELKF 716
 Clostridium/1-874

 SCLAV_5441/1-609 742 VRDEAEQVIAEVERATGVGLRLALGTMIELPRAALTAQIAEAAEFFSFCTNDLTQTVMGFSRDDVEASFETAYLEKGI FGVSPPFETIDRDGVGSLVRSAAA 844
 DSM738/1-905 742 VRDEAEQVIAEVERATGVGLRLALGTMIELPRAALTAQIAEAAEFFSFCTNDLTQTVMGFSRDDVEASFETAYLEKGI FGVSPPFETIDRDGVGSLVRSAAA 844
 GSK2/1-905 742 VRDEAEQVIAEVERATGVGLRLALGTMIELPRAALTAQIAEAAEFFSFCTNDLTQTVMGFSRDDVEASFETAYLEKGI FGVSPPFETIDRDGVGSLVRSAAA 844
 GSK6/1-905 734 VREEVERVLAESDTS GVPVTPCPVGTMIELPRAALTACGIAEAAEFFSFCTNDLTQTVMGFSRDDVEASFETAYLEKGI FGVSPPFETIDRDGVGSLVRSAAA 836
 SCO0208/1-898 746 VREEADQVIAEVAATGAKLKLISIGTMIELPRAALTACGIAEAAEFFSFCTNDLTQTVMGFSRDDVEASFETAYLEKGI FGVSPPFETIDKDGVGSLVAAA 848
 SCO2494/1-909 722 QVTILRQVAEKVFANVGKTI GYKVGVTMIELPRAALVADEIAEAAEFFSFCTNDLTQMTFGYSRDDVG--KFIPVYLAQGI LQHDPFEVLDRQGVGLVKFATER 823
 Zea/1-876 742 IREEVVKTAEAIVTKSGKRVHYTVGTMIELPRAAVTADSI AQKADFFSFCTNDLTQMCGGFSRDDAG--PFLRHYGNLGIYAQDPFQSIDQEGIGELVR IAVTK 843
 Trypanosoma/1-913 717 VKDVVVEVAEQVKEKSGDMQHYIGTMIELPRAALTADAIEAAEFFSFCTNDLTQMTFGFSRDDAG--KFLDSYKAKIYESDPFARLDQTGVGLVEMAVKK 818
 Clostridium/1-874

 SCLAV_5441/1-609 845 GRATRPGKLKLVGEHGGDPESVHFFHEVGLDYVSCSPFRIPVARLEAGRAALLSRHGGGR----- 905
 DSM738/1-905 845 GRATRPGKLKLVGEHGGDPESVHFFHEVGLDYVSCSPFRIPVARLEAGRAALLSRHGGGR----- 905
 GSK2/1-905 845 GRATRPGKLKLVGEHGGDPESVHFFHEVGLDYVSCSPFRIPVARLEAGRAALLSRHGGGR----- 905
 GSK6/1-905 837 GRATRPLDTIGVGEHGGDPDSVHFFHESAGLDYVSCSPFRVPVARLEAGRAALLADAGDSR----- 898
 SCO0208/1-898 849 GRATRPLDKLVGEHGGDPESVHFFHEVGLDYVSCSPFRIPVARLEAGRAASRSKGSDDR----- 909
 SCO2494/1-909 824 GRKARPNLKVIGGEHGGEPSSVAFFAKAGLDYVSCSPFRVP IARLAAAQVLV----- 876
 Zea/1-876 844 GRVKPMLKMGIGEHGGDPATIGFCHKVGLDYVSCSPFRVPVIAAAHASIKDRAAMKARKGFAAKL 913
 Trypanosoma/1-913 819 GRQTRPGKLKGVGEHGGDPSSVEFCHKVGLNYVSCSPFRVP IARLAAAQALNNK----- 874
 Clostridium/1-874

Figure 5.2 Multiple sequence alignment of PPKD amino acid sequences from *S. clavuligerus*, and *S. coelicolor*, *Z. mays*, *C. symbiosum*, and *T. brucei*. These amino acid sequences encode PPKD proteins used in this study and *T. brucei* [164], *Z. mays* [165], and *C. symbiosum* [157] have also had their crystal structures resolved. No catalytically active residues appear in the first half on the amino acid sequence, residues conferring active sites and binding sites are highlighted in a red and blue boxes respectively.

This alignment also contains a second PPDK identified from *S. clavuligerus* (SLCAV5441). This was excluded from the knockdown and over-expression experiments because, as illustrated in the MSA in Figure 5.2, this gene is two thirds the size of the other *Streptomyces ppdk*s suggesting that either the gene has been misannotated or the protein encoded may be truncated and non-functional. The MSA and the knowledge of the binding and active sites also supports this, as the truncated PPDK from *S. clavuligerus* (SCLAV5441) only appears to possess the Arg111 and His514 residues which correspond to one binding site and one of the two active sites respectively [158]. However, this protein is missing all other residues relating to binding sites of the protein and also the Cys residue corresponding with the second active site of the protein. Therefore, it was decided that this protein would not be pursued in this study.

5.5 Alignment shows PPK structures are conserved throughout different kingdoms of life.

Currently only three PPK proteins have their crystal structures resolved and, although distantly related, the sequences are well conserved. As the protein sequences are so well conserved it is expected that the corresponding 3-dimensional structures would be similar, however this isn't necessarily the case. Figure 5.3 shows the overlayed structures of the three proteins from *Z. mays* (green) [157], *C. symbiosum* (yellow) [149], and *T. brucei*. (blue) [156]. The structures were downloaded from UniProt and overlayed in PyMOL.

Each of the proteins has two active sites His458, Cys836, His455, Cys831, His483, and Cys856 for *Z. mays*, *C. symbiosum*, and *T. brucei* respectively which are shown in red and the binding sites shown in blue.

The binding site at the N-terminus of the protein is thought to be responsible for binding ATP and the binding site at the C-terminus of the protein is thought to be responsible for binding PEP/pyruvate. The central domain is thought to be particularly important as this is the location of the His residue which is thought to be responsible for shuttling the phosphate group between the two binding domains [157]. It is possible that any structural differences between these proteins could result in differences in functionality.

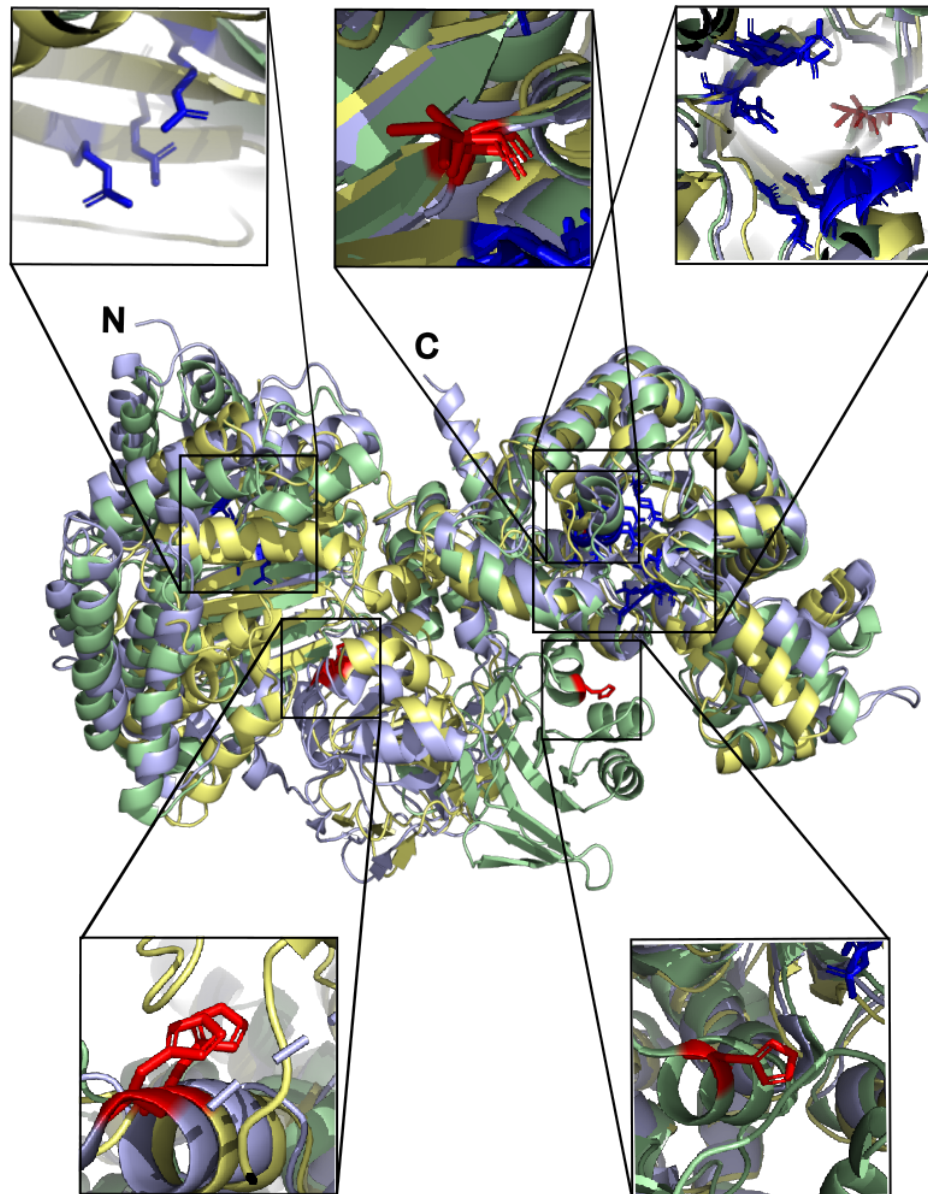


Figure 5.3 Overlaid structures of previously characterised PPDK proteins from *Z. mays*, *C. symbiosum*, and *T. brucei* shown in green, yellow, and blue respectively. PPDK proteins of *Z. mays* (UniProt: [P11155](#)), *C. symbiosum* (UniProt: [P22983](#)), and *T. brucei* (UniProt: [O76283](#)), were overlaid in PyMOL to show similarity of the structures. The active sites and binding sites of each are shown in red

and blue respectively. Figures were created in PyMOL using PDB files 1VBG, 2DIK, and 2X0S for *Z. mays*, *C. symbiosum*, and *T. brucei* respectively.

The root mean square deviation (RMSD) is a common method of quantifying similarity between two or more protein structures. RMSD is a measure of the distance between equivalent C α positions of specific residues or of the whole protein. A reasonable RMSD alignment score is < 2.5 – 3.0 Å [159]. The protein structures of *Z. mays* [154] and *T. brucei* [155] were both aligned to that of *C. symbiosum* [153] giving RMSD scores of 4.3 Å and 2.7 Å respectively. The alignment between *T. brucei* and *C. symbiosum* therefore suggests that they are similar in structure, much more than that of *Z. mays* and *C. symbiosum*. One reason that this could be the case is that eukaryotes are more closely related to bacteria than plants are. Another reason could be that in Maize the β -sheet containing one of the active sites seems to be oriented towards the C-terminus rather than oriented towards the N-terminus as it is in the bacterial and eukaryotic proteins.

5.6 Structures of PPK proteins from *S. clavuligerus* and *S. coelicolor* were predicted using AlphaFold2.

As a preliminary investigation, the structures of the PPK proteins from *S. coelicolor* (SCO0208 and SCO2494) and *S. clavuligerus* (SCLAV1689 and SCLAV5441) were predicted using the AlphaFold2 colab [160]. This exercise was to try to predict if these proteins were likely to be similar in structure to those of other organisms and, whilst no replacement for structural experiments such as X-ray crystallography or cryogenic Electron Microscopy (cryo EM), it is an accessible way to predict structures. The PDB files produced by AlphaFold2 were visualised in PyMOL with the confidence of the predicted structures coloured corresponding to the pLDDT score of each residue: magenta denotes a confidence score of 50 and cyan a confidence score of 100.

Predicted local distance difference test (pLDDT) is a method of scoring the confidence of a predicted structure by AlphaFold2 by assessing the disorder of regions [161]. A disordered region is of low confidence and has a pLDDT score of < 50 and a score of >70 is prediction of an area classed as a high confidence region [162]. It is anticipated from the sequence alignments that the structures will be somewhat similar to those PPK proteins already resolved, mainly in the areas of conserved sequence around active and binding sites, but it is expected that the *Streptomyces* PPKs will be almost identical to each other as they had 100% sequence identities.

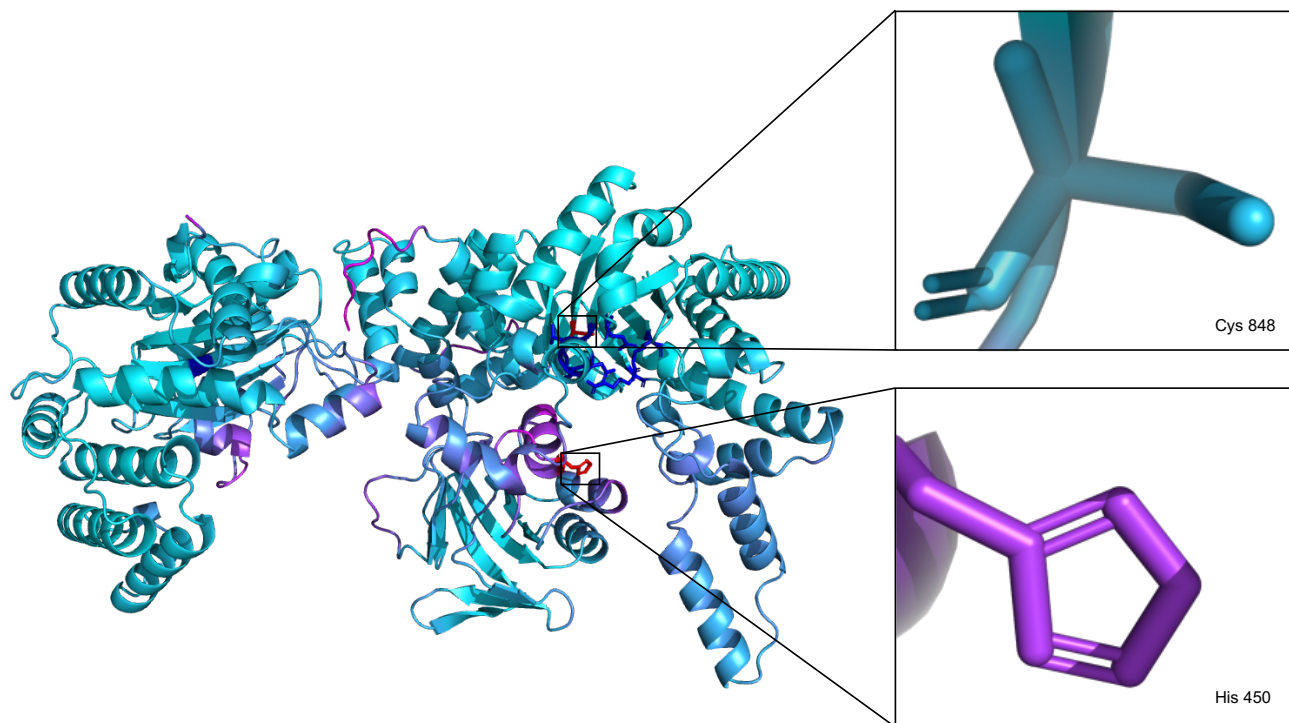


Figure 5.4 Structure of *S. coelicolor* PPK (SCO0208) as predicted by AlphaFold2 [160]. Predicted structure is coloured corresponding to the pLDDT score of each residue, magenta is equal to a confidence score of 50 and cyan a confidence score of 100. Binding sites and active sites shown in blue and red respectively.

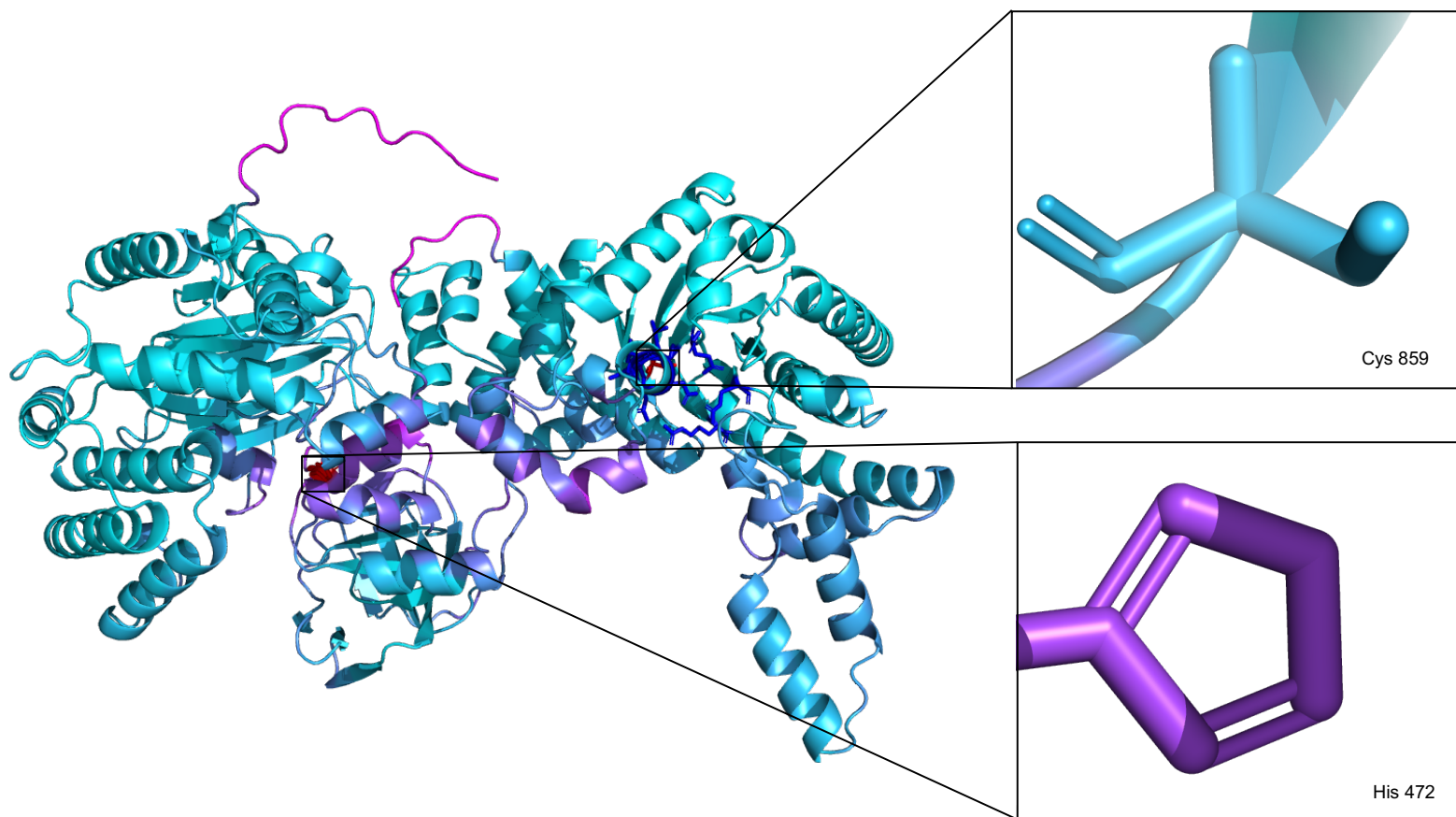


Figure 5.5 Structure of *S. coelicolor* PPDk (SC2494) as predicted by AlphaFold2 [160]. Predicted structures are coloured corresponding to the pLDDT score of each residue, magenta is equal to a confidence score of 50 and cyan a confidence score of 100. Binding sites and active sites shown in blue and red respectively.

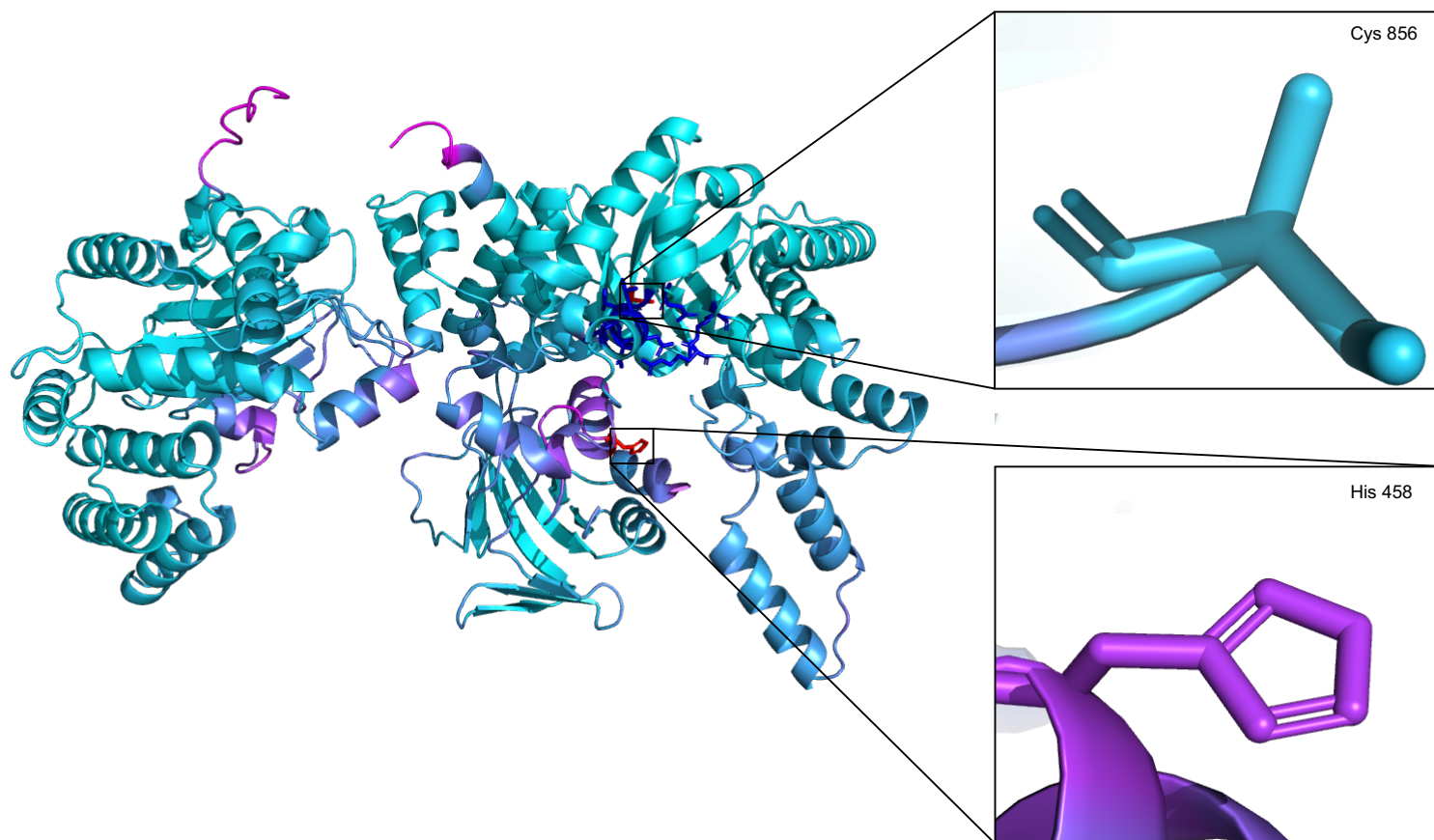


Figure 5.6 Structure of *S. clavuligerus* PPDK (SCLAV1689) as predicted by AlphaFold2 [160]. Predicted structures are coloured corresponding to the pLDDT score of each residue, magenta is equal to a confidence score of 50 and cyan a confidence score of 100. Binding sites and active sites shown in blue and red respectively.

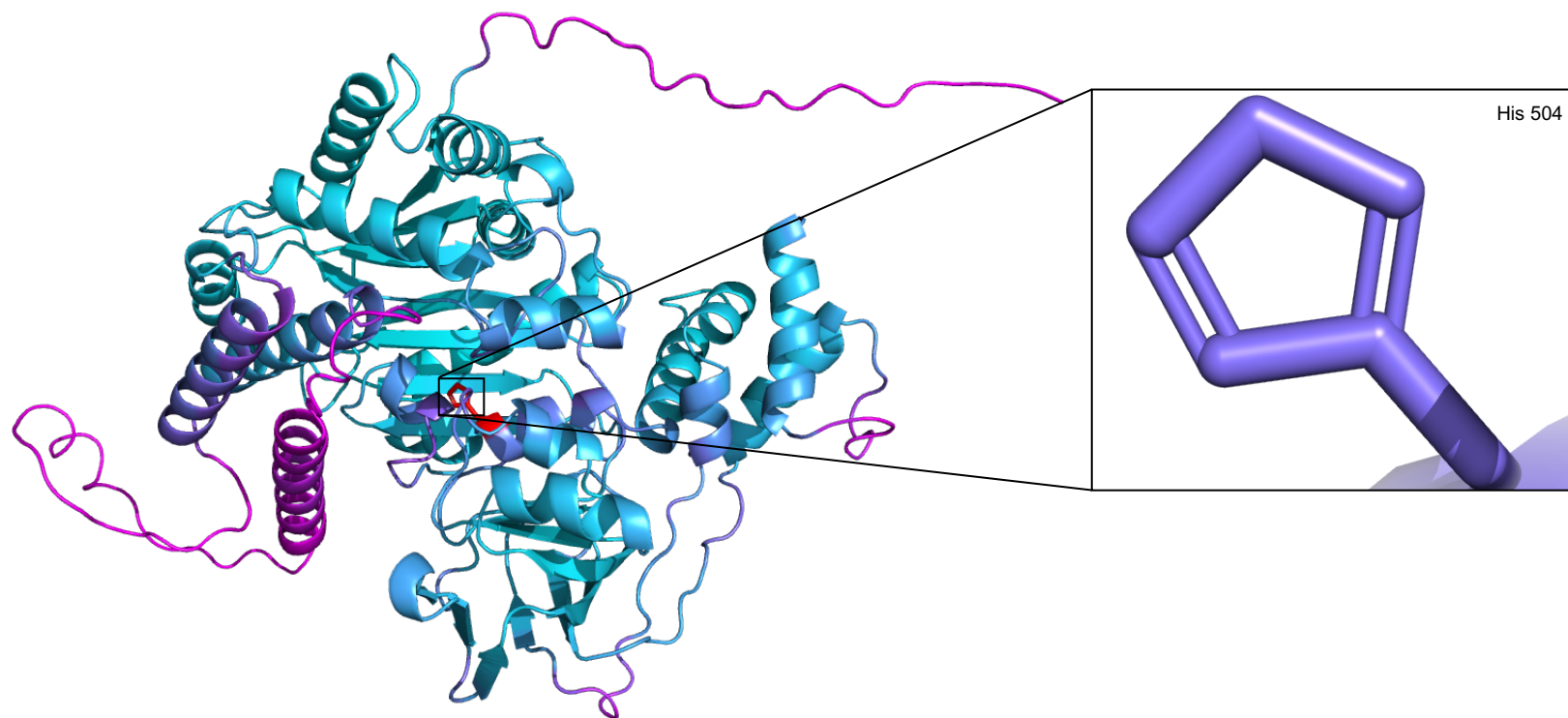


Figure 5.7 Structure of truncated *S. clavuligerus* PPDK (SCLAV5441) as predicted by AlphaFold2 [160]. Predicted structures are coloured corresponding to the pLDDT score of each residue, magenta is equal to a confidence score of 50 and cyan a confidence score of 100. Binding sites and active sites shown in blue and red respectively.

The pLDDT scores for each of these proteins is high although unsurprisingly the confidence scores around the N and C termini are lower than the rest of the protein. The proteins also appear to have a lower confidence score around the central domain of the protein where there is some variation in the architecture. Although the PPDK sequences in *S. coelicolor* and *S. clavuligerus* are highly conserved, one particularly notable difference between the predicted structures of SCO2494 PPDK2 and its orthologues SCO0208 PPDK1 and SCLAV1689 PPDK is the orientation of the β -sheet in the central domain of the protein. In the predicted structure of *S. coelicolor* PPDK2 the β -sheet appears to be oriented towards the N terminus of the protein while in the predicted structures of *S. clavuligerus* PPDK (SCLAV1689) and *S. coelicolor* PPDK1 (SCO0208) the β sheet appears to be oriented towards the C terminus of the protein. The N terminal domain binds PEP/PYR [157, 158].

5.7 Construction of vectors for the overexpression of Streptomycete PPDK proteins.

Vectors had to be constructed to enable overexpression of the proteins under study. The sequences for the two *S. coelicolor* genes (SCO0208 and SCO2494) and the single *S. clavuligerus* gene (SCLAV1689) encoding the PPDK proteins were downloaded from the StrepDB database (<https://strepdb.streptomyces.org.uk/>). These sequences were codon optimised for heterologous expression in *E. coli*, as the high GC content and codon usage in *Streptomyces* often limits the expression of proteins in *E. coli*. These sequences were then synthesised by Genescript. The sequences for each gene were subcloned into the pET-14b expression vector using *Nde*I and *Xho*I restriction sites. Figure 5.8 shows the maps for each of the three expression vectors. The *ppdk* genes are expressed under the T7 promoter and in frame with a N-terminal 6xHis tag which is upstream of a thrombin cleavage site.

Because the structures of these proteins, predicted by AlphaFold2, suggested that neither the N nor C terminus of the protein is occluded once folded, the decision was made to tag the N terminus of the protein with a 6xHis tag to facilitate purification.

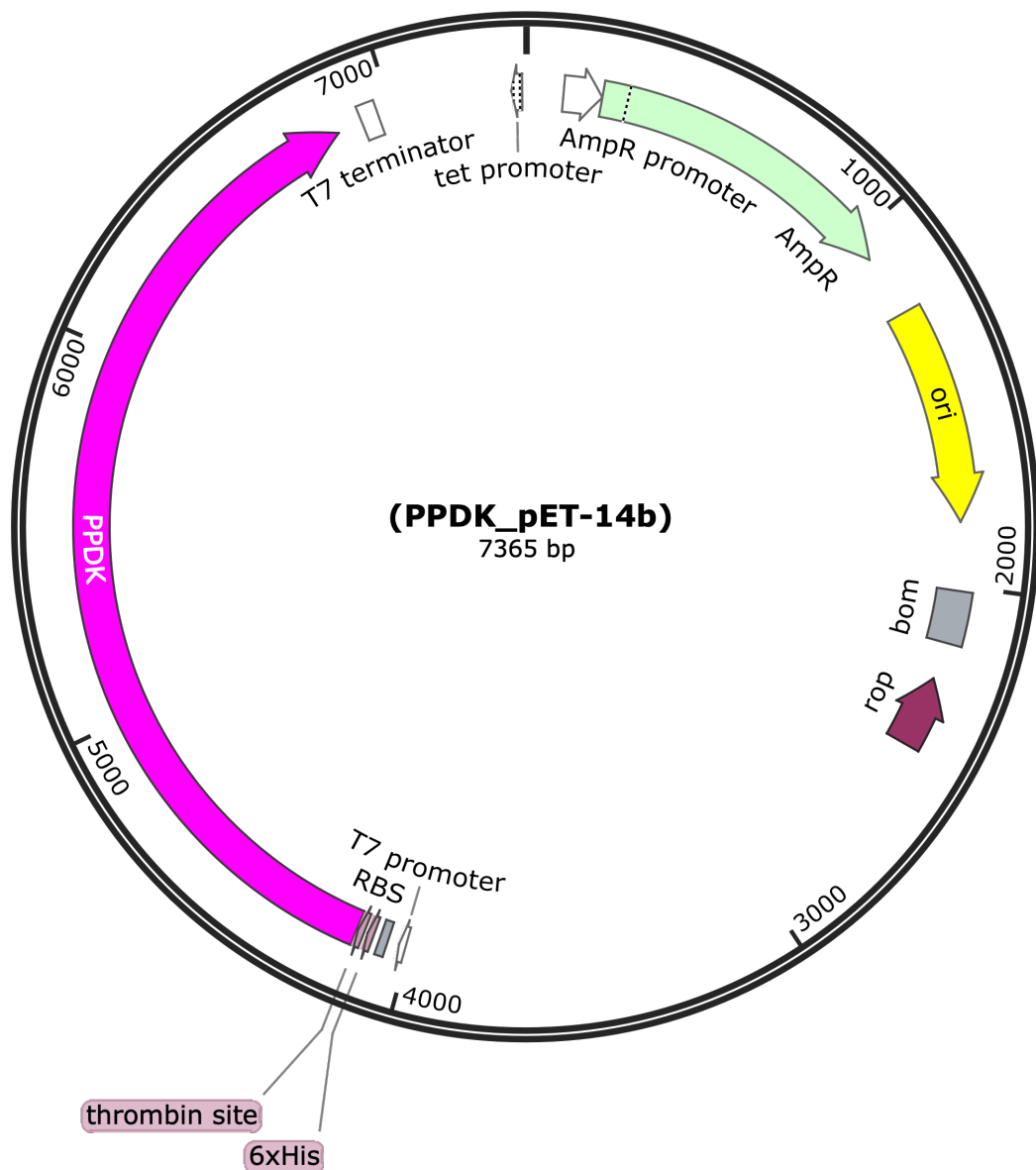


Figure 5.8 Schematic representation of overexpression constructs for PPDK from *Streptomyces*. Each vector was built using a pET14b backbone, with the CDS of the relevant gene sequence inserted downstream of the T7 promoter. The gene sequences were modified to add an N-terminal hexa-his-tag to aid in purification of the gene product.

5.8 Overexpression of Streptomycete PPK proteins

To determine the conditions required for optimal overexpression of the PPK proteins, *E. coli* BL21 (DE3) was transformed with each of the plasmids encoding the His-tagged *ppk* genes from the *Streptomyces* species.

Due to the limited information about these proteins the first expression trials were carried out at three different temperatures, 18 °C, 26 °C, and 37 °C, giving a wide range of temperature conditions for the expression. These temperatures were chosen because 18 °C is the temperature at which *Streptomyces* Pyk proteins are expressed. 26 °C and 37 °C because these are the temperatures used to grow *S. clavuligerus* and *E. coli* respectively. For each protein the whole cell lysates of the uninduced culture (U) and the induced culture at each temperature were visualised on an SDS-PAGE gel (Figure 5.9).

It was expected that if any of these conditions were suitable for the expression of the PPK proteins a band at ~ 96 kDa would be visible on the SDS – PAGE gel.

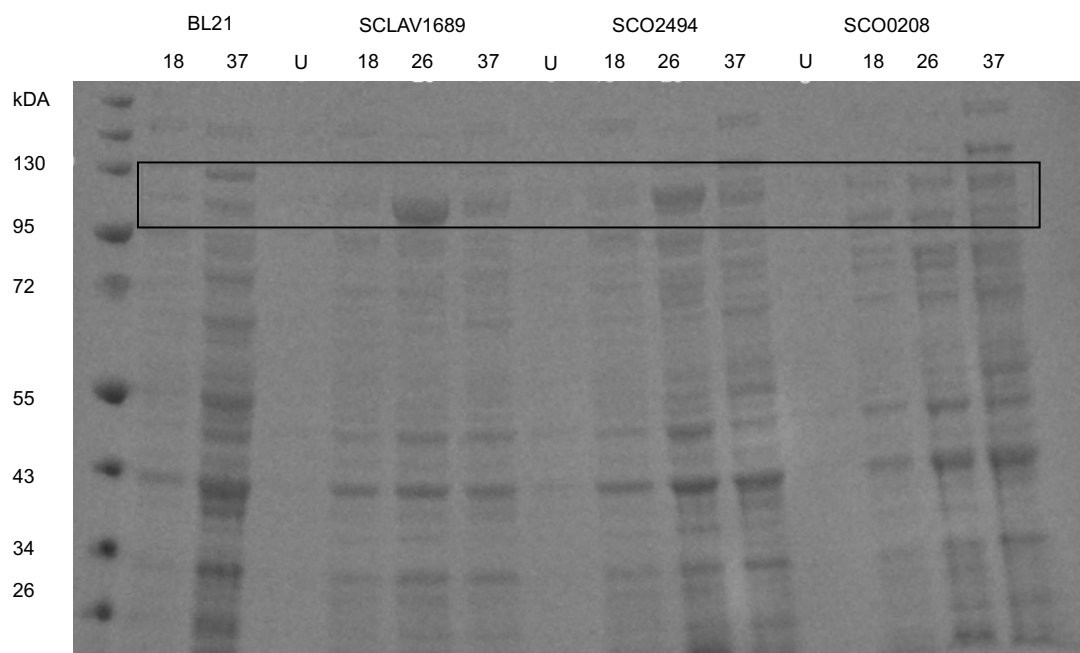


Figure 5.9. Expression of the PPK proteins from *Streptomyces* at 18 °C, 26 °C, and 37 °C. Cultures were grown at three different temperatures, 18 °C, 26 °C and 37 °C. For each protein the whole cell lysate of the uninduced culture (U) and the induced culture at each temperature were examined by SDS-PAGE gel electrophoresis. Color Prestained Protein Standard, Broad Range. 10-250 kDa by New England Biolabs was used as a ladder.

The PPDK proteins are visible at ~ 96 KDa with the uninduced *E. coli* BL21 as a negative control closest to the marker in Figure 5.9. While bands of ~96 kDa were observed for PPDK SCO0208, they were similar to bands present in the uninduced control and did not vary with temperature, suggesting either no or low amounts of protein had been expressed. In contrast, strong bands ~96 kDa were observed for both PPDK 1689 and PPDK 2494 when expressed at 26 °C. This suggests that both proteins express well at 26 °C, and thus this temperature was selected as the optimal expression temperature for purification trials.

5.9 Optimisation of PPDK purification

Once it had been determined that it was possible to express the PPDK proteins, purification trials could commence using IMAC. Crude lysates obtained from cell pellets from the over-expression trials were passed through a 1 ml His Trap FF Crude column (GE Healthcare) loaded with NiSO₄, after being equilibrated with buffer A, by hand using a 20 ml syringe. Following binding of the proteins, the column was washed with IMAC buffer A and bound proteins were eluted in five column volumes of IMAC buffer B containing 500 mM imidazole (see Table 2.5), which is the highest concentration of imidazole that would be used and should remove any protein bound to the column.

Figure 5.10 shows that a band is present for *S. coelicolor* PPDK 1 (SCO0208) at temperatures 26 °C and 37 °C (Figure 5.10 A) indicating that those fractions contained the protein of interest. The band for SCO0208 at 26 °C is slightly darker than that at 37 °C suggesting that expressing at 26 °C yields slightly more protein.

The *S. clavuligerus* PPDK (SCLAV1689) also shows a band from the 26 °C culture, indicating that PPDK can be purified from cultures grown at 26 °C (Figure 5.10 B).

The *S. coelicolor* PPDK 2 (SCO2494) can also be purified from cultures grown at 26 °C as illustrated by the bands at ~96 kDa in Figure 5.10 B.

As protein from all three expression cultures grown at 26 °C using autoinduction media could be purified using IMAC, these conditions were used for expression of all three proteins moving forward to larger scale purifications.

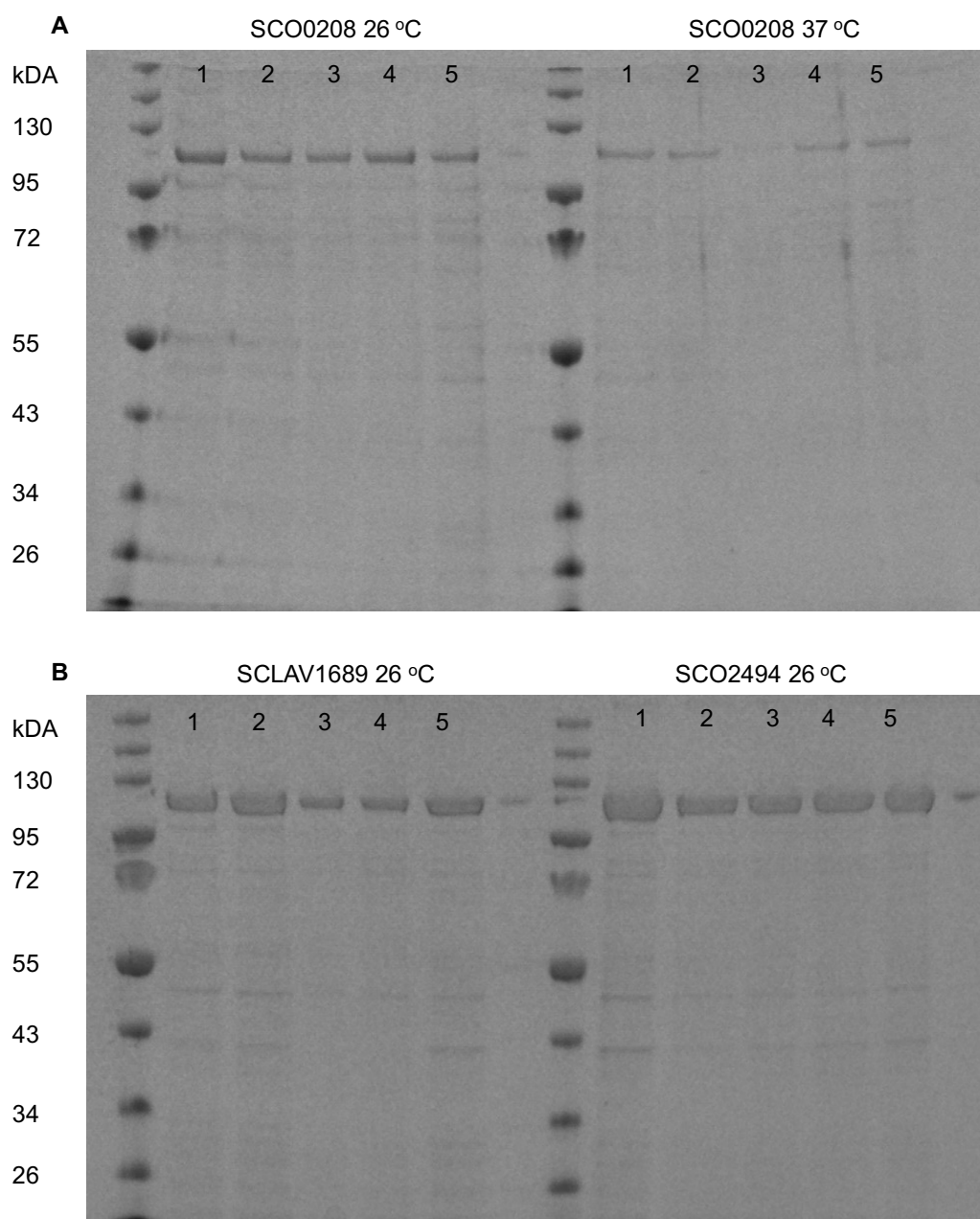


Figure 5.10 Crude purification of *Streptomyces* PPDK proteins by hand-loaded His Trap FF Crude columns. PPDK proteins are shown by bands on the SDS-PAGE gel at ~96 kDa. A. *S. coelicolor* PPDK1 (SCO0208) expressed at 26 °C (left) and 37 °C (right). B. *S. clavuligerus* PPDK (SCLAV1689; left) and the *S. coelicolor* PPDK2 (SCO2494; right) expressed at 26 °C. The SDS-PAGE gels show samples of each 1 ml fraction eluted from the IMAC column using buffer B containing 500 mM of imidazole.

5.9.1 Gradient elution unexpectedly reduced yield of protein

The crude purification showed that protein expressed at 26 °C could be subsequently purified using a nickel charged 1 ml His trap column. In addition, a single-step elution using 500 mM imidazole liberated a significant amount of protein from the column. This was sufficient to confirm that the His-tag on all three proteins was accessible and that IMAC purification was viable. However, it was necessary to scale-up and further optimise the purification to produce sufficient protein for functional characterisation.

To optimise the purification, the expression cultures were scaled up to 0.5 L cultures, using a 5% inoculum from overnight cultures of *E. coli* BL21 DE3 in autoinduction media (Table 2.4). The resulting cells were collected by centrifugation (Beckman High Speed Centrifuge, JLA-9.1000 at 7,439 x *g* for 45 mins) and resuspended in IMAC Buffer A before being lysed by sonication. The cells were sonicated on ice in 10 bursts of 10 seconds with 50 second intervals. The supernatant was then clarified by centrifugation (Beckman High Speed Centrifuge, JA-25.50 at 27,220 x *g* for 30 mins). Although the trial IMAC purification was a step elution from 0 mM to 500 mM, imidazole can cause damage to the proteins which means it is best to find the minimal concentration required to elute the proteins. Therefore, the first round of purification using the AKTA (GE Healthcare) employed a gradient elution from 8% (40 mM) imidazole to 100% (500 mM) imidazole IMAC buffer B. This should result in a peak on the chromatogram at the lowest concentration of imidazole required to elute the protein.

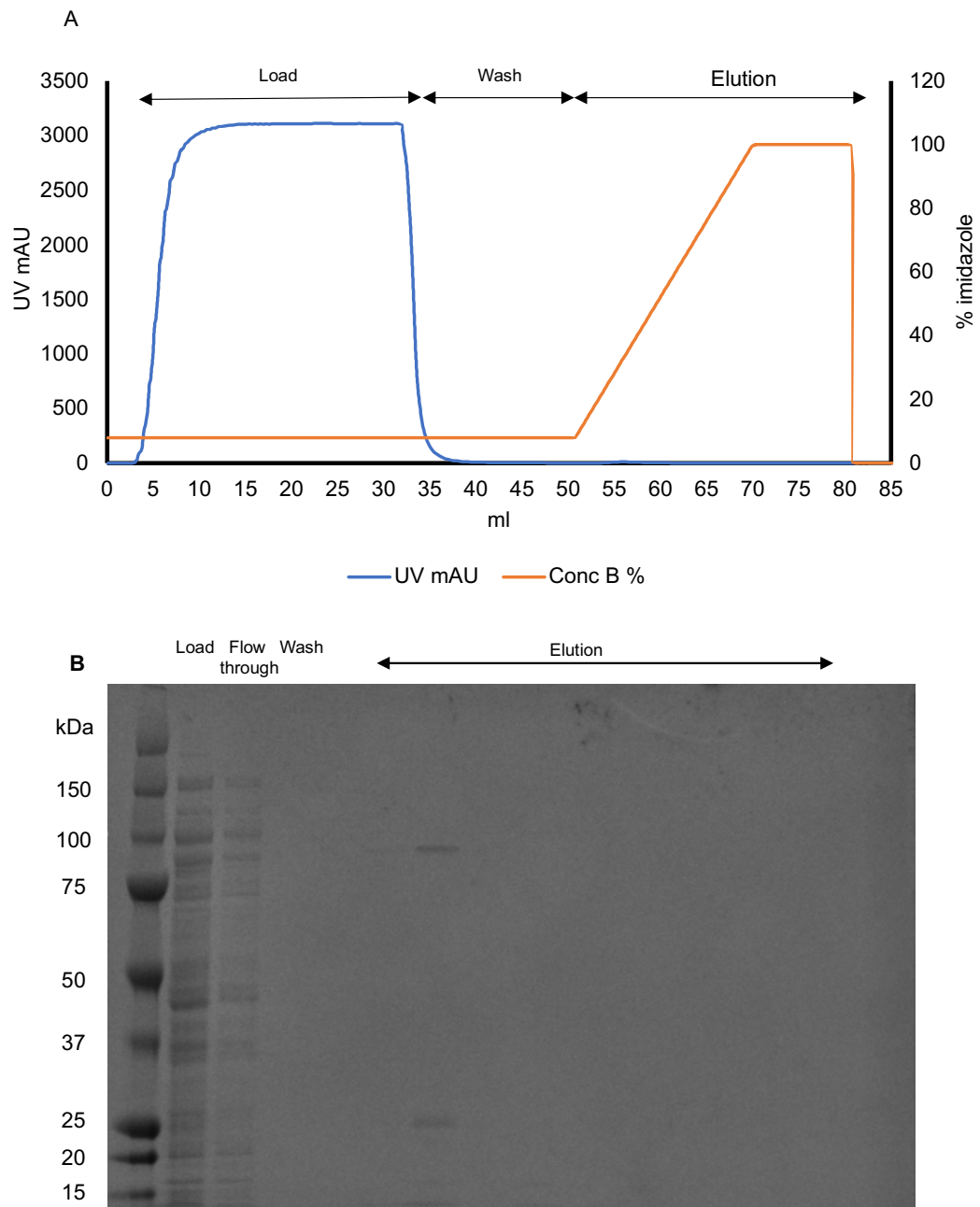


Figure 5.11 The UV chromatogram and SDS-PAGE analysis of gradient purification of *S. clavuligerus* PPDK (SCLAV1689). A. Chromatogram of the gradient elution of *S. clavuligerus* PPDK (SCLAV1689). B. SDS-PAGE gel shows bands ~96 kDa in the load, flow through, and elution.

The chromatogram (Figure 5.11) of the gradient purification of *S. clavuligerus* PPDK (SCLAV1689) using the AKTA shows an increase in UV during the load step however there was no UV peak appearing on the chromatogram 5.11 A during the elution step, which would indicate protein was being loaded onto the column but then not being eluted off the column. Analysis of the fractions by SDS-PAGE shows that PPDK protein was present in the load, as there is a band ~ 96 kDa. However, this was also present in the flow through indicating that the protein was not binding to the column properly. The ~96 kDa band present in the elution represents the PPDK protein showing that some successfully bound and was subsequently eluted from the column in the third millilitre of the elution which is at an approximate concentration of 200 mM imidazole.

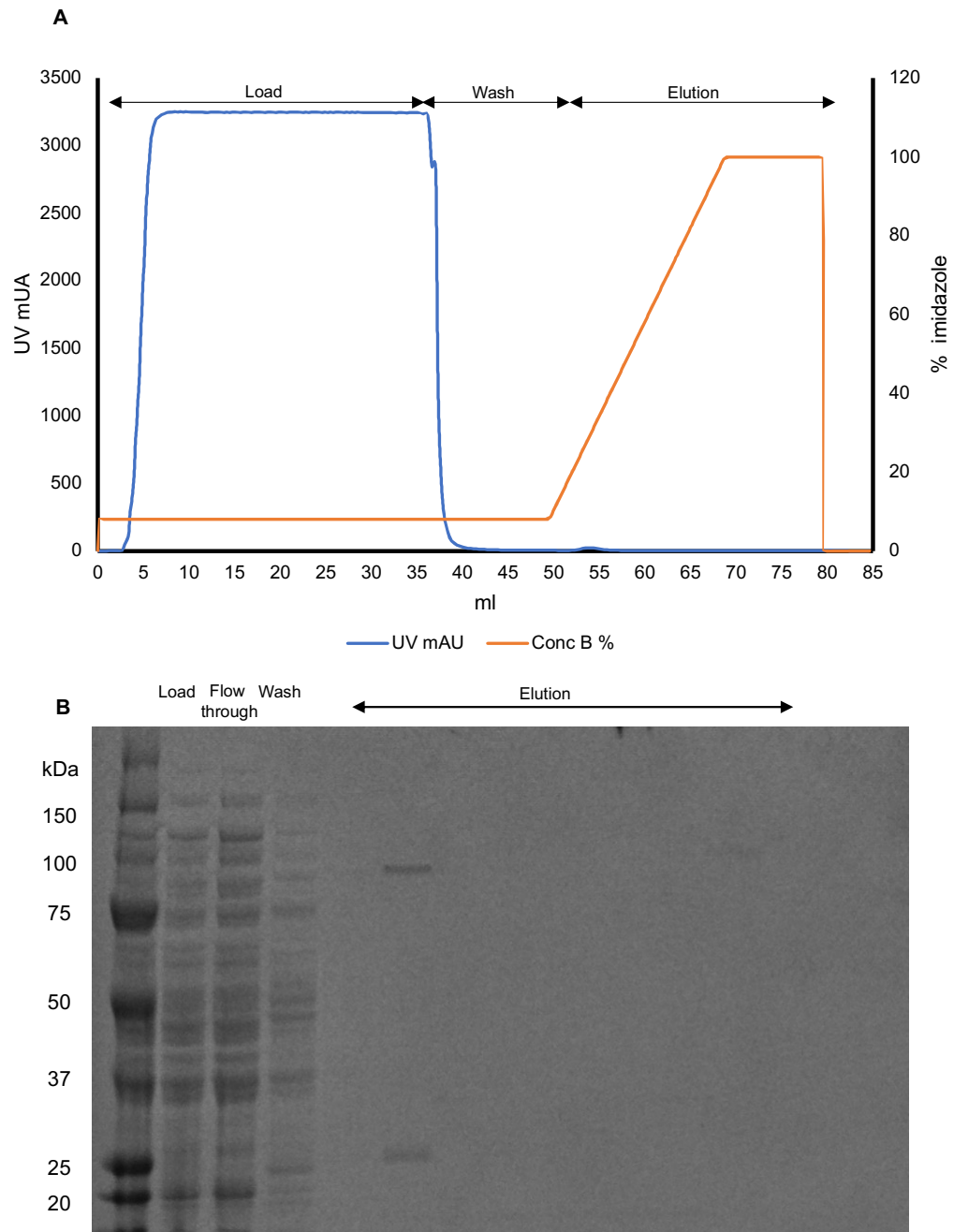


Figure 5.12 The UV chromatogram and SDS-PAGE analysis of gradient purification of *S. coelicolor* PPK (SCO2494). A. Chromatogram of the gradient elution of *S. coelicolor* PPK2 (SCO2494). B. SDS-PAGE gel shows bands ~98 kDa in the load, flow through, wash, and elution.

The chromatogram (Figure 5.12 A) of the gradient purification of *S. coelicolor* PPDK2 (SCO2494) using the AKTA shows an increase in UV during the load step and a very small UV peak appearing on the chromatogram during the elution step. This indicates protein was loaded onto the column and suggested a small amount of protein was being eluted from the column. Analysis of the fractions by SDS-PAGE shows that PPDK protein was present in the load, as there is a band ~ 98 kDa. However, this was also present in the flow through and wash indicating that the protein was not binding to the column properly, which could be due to overloading. The ~98 kDa band present in the elution represents the ppdK protein showing that PPDK2 was successfully bound and subsequently eluted from the column in the first four millilitres of the elution, which corresponds to range of approximately 8-200 mM imidazole concentration.

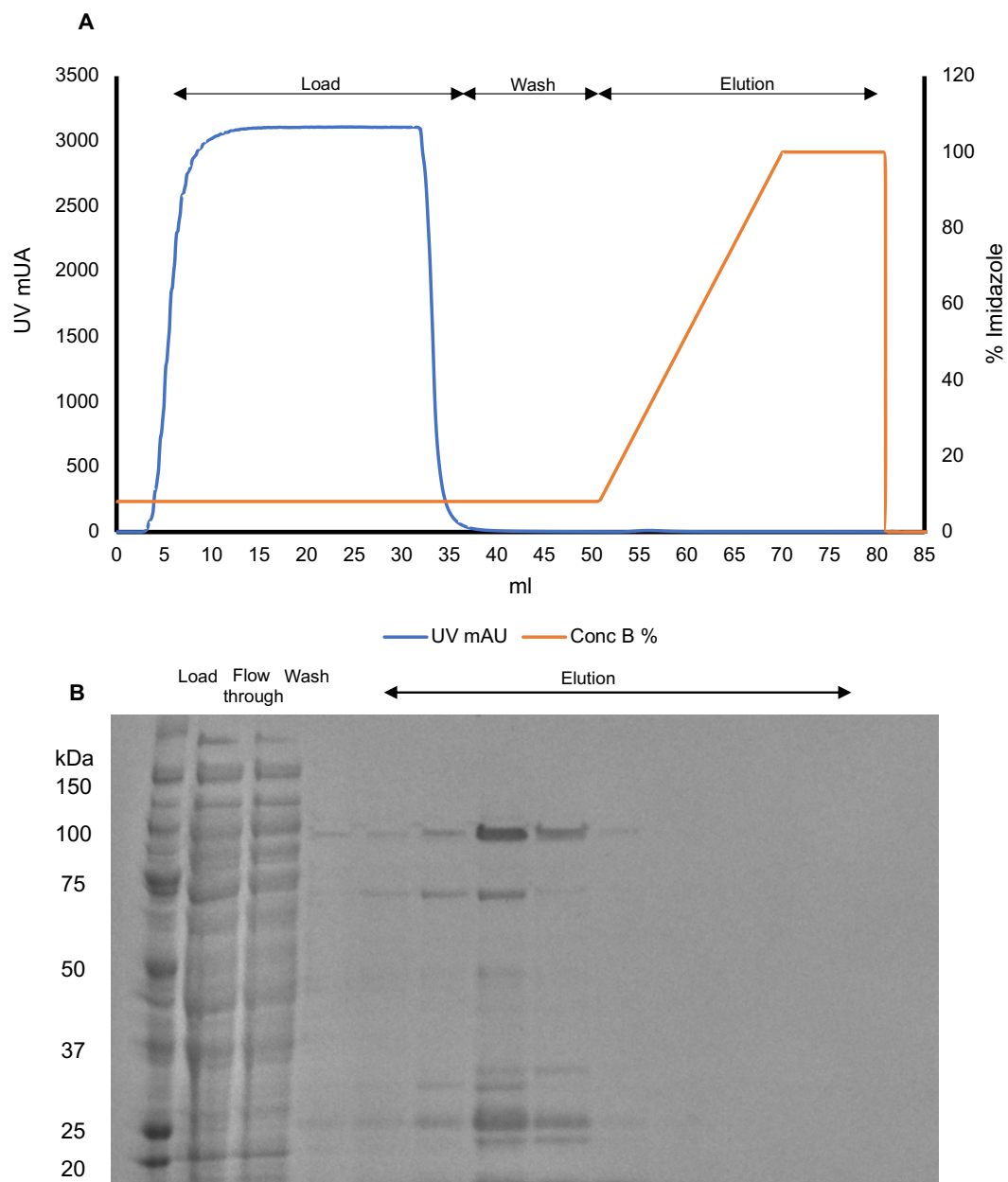


Figure 5.13 The UV chromatogram and SDS-PAGE analysis of gradient purification of *S. coelicolor* PPDK 1 (SCO0208). A. Chromatogram of the gradient elution of *S. coelicolor* PPDK1 (SCO0208). B. SDS-PAGE gel shows bands ~98 kDa in the load, flow through, wash, and elution.

The chromatogram (Figure 5.13 A) of the gradient purification of *S. coelicolor* PPK1 (SCO0208) using the AKTA shows an increase in UV during the load step and a very small UV peak appearing on the chromatogram during the elution step. This indicated protein was loaded onto the column and suggested a small amount of protein was being eluted from the column. Analysis of the fractions by SDS-PAGE shows that PPK protein was present in the load, as there is a band ~ 98 kDa. However, this was also present in the flow through and wash indicating that the protein was not binding to the column properly. The ~98 kDa band present in the elution represents the PPK protein showing that PPK1 (SCO0208) was successfully bound and subsequently eluted from the column in ml 2 of the gradient elution (in the range of approximately 8-200 mM imidazole concentration).

Although it was possible to purify the proteins previously when expressed at 26 °C using the combination of IMAC and a step elution protocol, when the cultures were scaled up and a gradient elution was employed using AKTA very little protein could be purified. Although the elution had changed from a step elution to a gradient elution, if the protein was present and bound to the column this should still result in purification of the protein as the final imidazole concentration of the gradient is equal to that of the step elution. This suggests that the proteins may no longer be being expressed properly. This could be due to the larger expression cultures, although still inoculated with a 5% seed culture as in the initial trials, may be taking longer to grow. If the cultures were not growing as quickly it will take longer for the glucose in the autoinduction media to be utilised. Autoinduction media works by inducer exclusion, meaning that *E. coli* is able to grow on the glucose before starting to express the recombinant protein which occurs once all the glucose has been utilised. At that point it will begin to utilise the second carbon source present, lactose which is what induces heterologous expression of the recombinant proteins. Lactose is transported into the

cells by *lac* permease for utilisation and a portion of this is converted to allolactose by β -galactosidase. Allolactose binds to the *lac* repressor enabling T7 RNA polymerase to bind to the T7 promotor under the control of which are the *ppdk* genes [163].

5.9.2 Purifying PPDK proteins through step elution increases yield of protein.

After attempting to purify the three PPDK proteins (SCLAV1689; SCO2492; SCO0208) and recovering only small amounts of protein, the *ppdk* genes were expressed again in 50 mL cultures at 26 °C as in the expression trials. The cultures were still inoculated 5% (v/v) from the seed culture *E. coli* BL21 DE3 in 50 mL of autoinduction media. After the cells had grown overnight, they were collected by centrifugation and lysed by sonication (as previously outlined). Purification was carried out using the AKTA (GE Healthcare) but with a step elution protocol, with an elution step of 60% (300 mM imidazole) and another at 100% (500 mM imidazole) of IMAC buffer B.

This replicated the purification conditions from the trial purification. Elution would be in two steps, at a concentration of 300 mM and 500 mM. 300 mM is the concentration at which a small peak can be seen on the chromatograms 5.12 A and 5.13 A these small peaks also correspond with the fraction in which protein is visible on the SDS-PAGE gels.

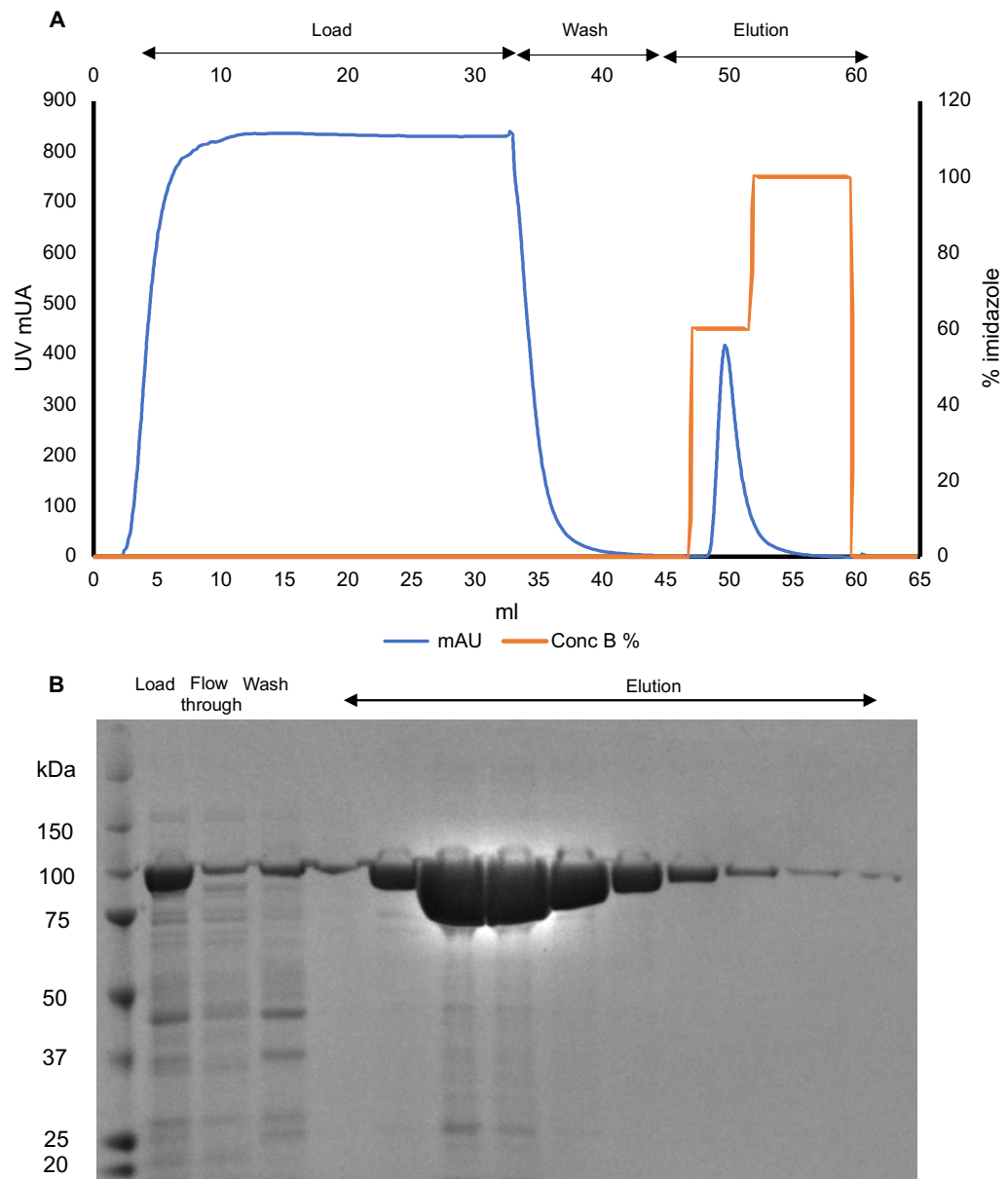


Figure 5.14 The UV chromatogram and SDS-PAGE analysis of step purification of *S. clavuligerus* PPDK (SCLAV1689). A. Chromatogram of the step elution of *S. coelicolor* *S. clavuligerus* PPDK (SCLAV1689). B. SDS-PAGE gel shows bands ~98 kDa in the load, flow through, wash, and elution.

The step purification of *S. clavuligerus* PPDK (SCLAV1689) using the AKTA shown in Figure 5.14 A shows an increase in UV during the load step and a peak of approximately 450 mUA appearing on the chromatogram during the 60% elution step. This shows that protein was loaded onto the column and protein was successfully eluted at a concentration of 300 mM of imidazole in IMAC buffer B. Analysis of the fractions by SDS-PAGE shows that PPDK protein was eluted as there are bands at ~98 kDa. It also shows a ~98 kDa band was present in the flow through and wash as well as the load suggesting that the 1 ml HisTrap column could have been overloaded. The ~98 kDa band present in the elution represents the PPDK protein showing that protein was successfully bound and subsequently eluted from the column over 10 millilitres concentrated in the 300 mM elution step.

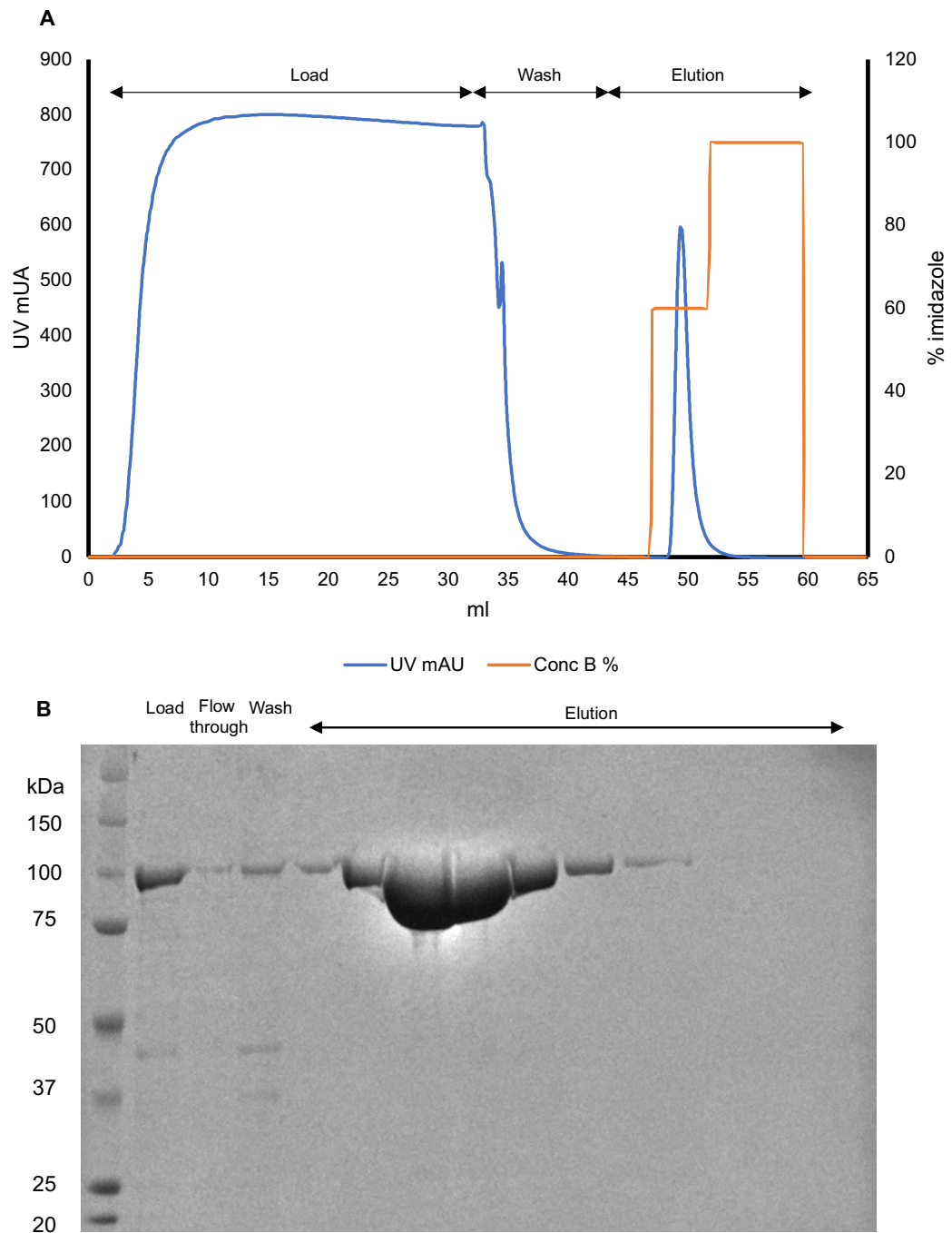


Figure 5.15 The UV chromatogram and SDS-PAGE analysis of step purification of *S. coelicolor* PPK2 (SCO2494). A. Chromatogram of the step elution of *S. coelicolor* PPK2 (SCO2494). B. SDS-PAGE gel shows bands ~98 kDa in the load, flow through, wash, and elution.

The step purification of *S. coelicolor* PPDK2 (SCO2494) using the AKTA shown in Figure 5.15 A shows an increase in UV during the load step and a UV peak appearing on the chromatogram during the first elution step of approximately 600 mUA. This shows that protein was loaded onto the column and successfully eluted from the column during the 300 mM elution step. Analysis of the fractions by SDS-PAGE shows that PPDK protein was present in the load, as there is a band ~ 96 kDa. However, there is also a faint band present in the flow through and wash potentially again indicating that the column was saturated. The ~96 kDa band present in the elution represents the PPDK protein showing that PPDK2 was successfully bound and subsequently eluted from the column during the first elution step where the concentration of imidazole in IMAC buffer B was 300 mM.

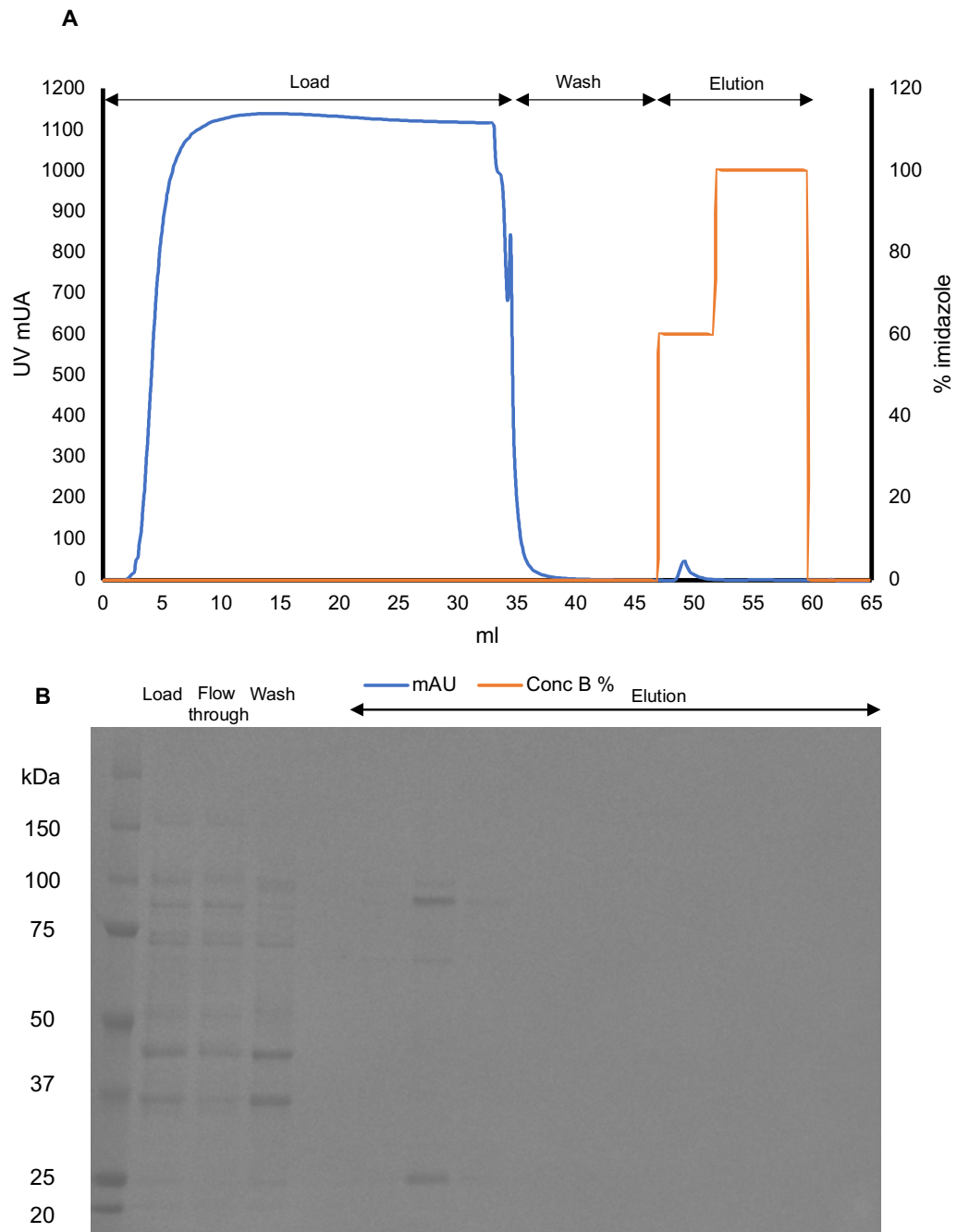


Figure 5.16 The UV chromatogram and SDS-PAGE analysis of step purification of *S. coelicolor* PPK1 (SCO0208). A. Chromatogram of the gradient elution of *S. coelicolor* PPK1 (SCO0208). B. SDS-PAGE gel shows bands ~98 kDa in the load, flow through, and elution.

The step purification of *S. coelicolor* PPK1 (SCO0208) using the AKTA shown in Figure 5.16 A shows an increase in UV during the load step and a small UV peak of approximately 50 mUA appearing on the chromatogram during the first elution step. This shows that protein was loaded onto the column and a small amount of protein was eluted from the column at a concentration of 300 mM. Analysis of the fractions by SDS-PAGE shows that PPK protein was present in the load, as there is a faint band ~ 98 kDa. However, this was also present in the flow through which could suggest that the protein was not binding well to the column. The ~98 kDa band present in the elution represents the PPK protein showing that PPK1 was successfully bound and subsequently eluted from the column during the first elution step, at a concentration of 300 mM imidazole IMAC buffer B.

5.10 Enzymatic characterisation of purified PPDK proteins.

The PPDK enzymes isolated from plants generally function in the direction of PEP formation, but in some parasites and bacteria there is evidence that the pyruvate forming direction is possible in the absence of pyruvate kinases [164]. Due to evidence in Chapter 4, it is important to try and understand the functionality of these enzymes. If they can function bi-directionally and what the kinetic constants are for each direction.

The PEP forming reaction catalysed by PPDK can be split into three stages as illustrated in Figure 5.17. Stage I) two phosphates from ATP are donated to the enzyme. Stage II) one of the two phosphates from the enzyme complex is donated to an orthophosphate molecule, resulting in the formation of a diphosphate molecule and an enzyme-phosphate complex. Stage III) the phosphate from the enzyme complex is transferred to the pyruvate resulting in free enzyme and phosphoenolpyruvate (PEP) [165].

The pyruvate forming reaction catalysed by PPDK uses PEP and AMP to yield pyruvate and ATP shown in Figure 5.18. To assay PPDK activity, this reaction is coupled with lactate dehydrogenase that converts the pyruvate produced by PPDK to lactate with oxidation of the cofactor NADH, which can be used to measure the rate of activity [40, 166].



Figure 5.17 The tripartite PPDK reaction in the phosphoenolpyruvate yielding direction.



Figure 5.18 The PPDk reaction in the pyruvate yielding direction coupled with the lactate forming LDH reaction. This reaction is often performed by pyruvate kinase in bacterial species, this coupled reaction is used to measure pyruvate formation.

PPDK activity in the gluconeogenic (PEP forming) direction, was assessed using a commercially available PPDK assay kit (Abcam). The kit employs a colorimetric assay, wherein intermediate products produced by PPDK react with a proprietary enzyme mix and a dye probe to produce a complex that can be quantified by measurement at 570 nm.

PPDK activity was measured for all three *Streptomyces* proteins (Figure 5.19). These data suggested that the proteins could be purified in an active state. The lowest specific activity was observed in SCO0208 and the orthologous protein to SCLAV1689. The highest specific activity was observed from the SCO2492 protein. To better characterise the proteins a kinetic study would be required for all enzymes in the PEP forming direction and the pyruvate forming direction to examine if these enzymes were potentially bifunctional under the correct physiological conditions. However, due to the lack of availability of the Abcam kit, it was not possible to perform full kinetic characterisation in the PEP forming direction.

It is important that a complete characterisation of the kinetics is conducted to elucidate the distinct functions of the isoenzymes in *S. coelicolor*. This would enable direct comparisons via parameters such as K_{cat} for these enzymes to explore rates of turnover in the two different species studied here. Kinetic characterisation was attempted but it was not possible to complete these experiments and draw any conclusions about the enzyme parameters.

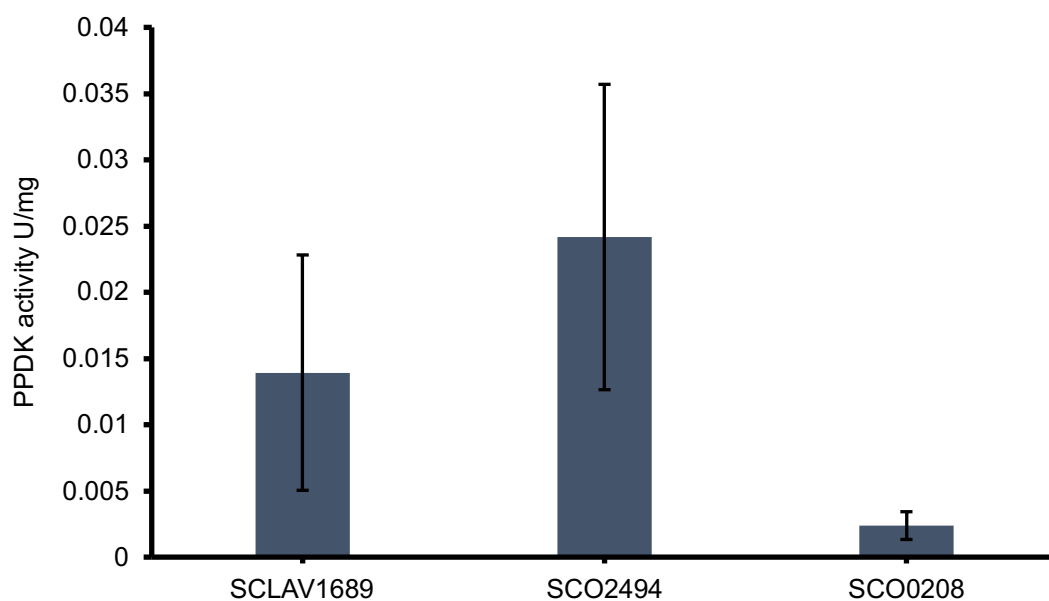


Figure 5.19 PEP-forming activity of *Streptomyces* PPDKs. Activity assay was carried out using the PPDK whole cell lysate activity assay kit supplied by Abcam. The assay was carried out in triplicate using purified proteins. Average activity (of three replicates) plotted and error bars show standard error of the mean.

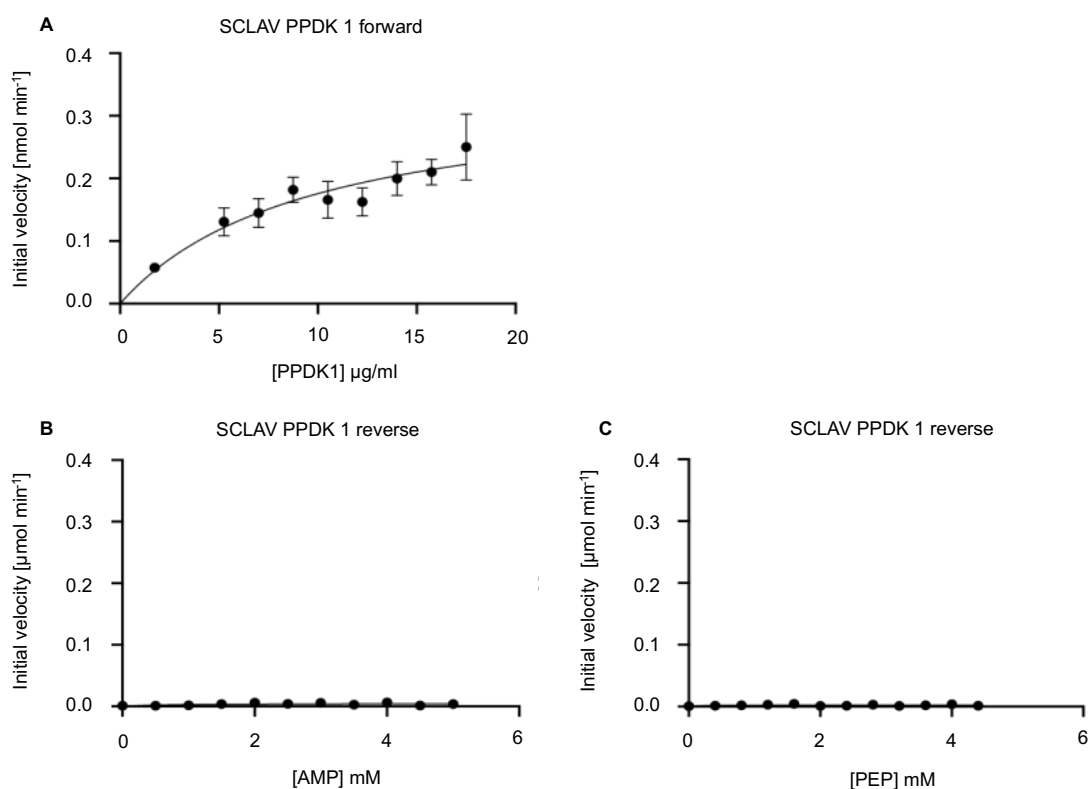


Figure 5.20 Enzyme kinetics experiments for *S. clavuligerus* PPDK (SCLAV1689) in both the PEP and PYR forming directions. A. data acquired when PPDK is performing forward PEP forming reaction. B and C show PPDK the reverse Pyruvate forming using ATP and PEP as the variable substrate respectively. When AMP was the rate limiting substrate a concentration of 10 mM PEP was used, and when PEP was the rate limiting substrate a concentration of 0.5 mM AMP was used.

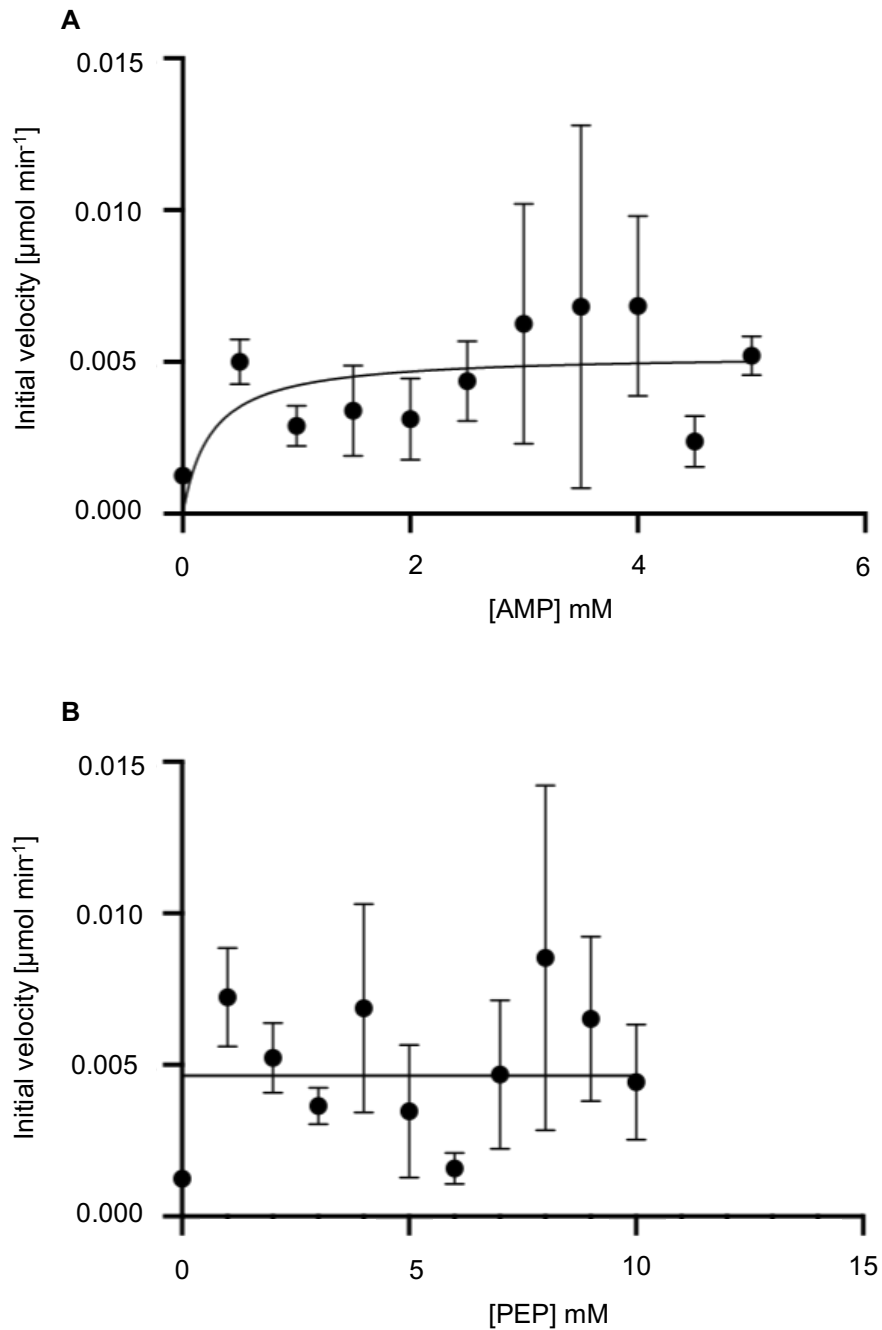


Figure 5.21 Enzyme kinetics experiments for *S. coelicolor* PPDk2 (SCO2494) in the reverse (pyruvate forming) direction. A and B show PPDk activity in the reverse (pyruvate forming reaction) using ATP and PEP as the variable substrate respectively. When AMP was the rate limiting substrate a concentration of 10 mM PEP was used, and when PEP was the rate limiting substrate a concentration of 0.5 mM AMP was used. These data show no trend.

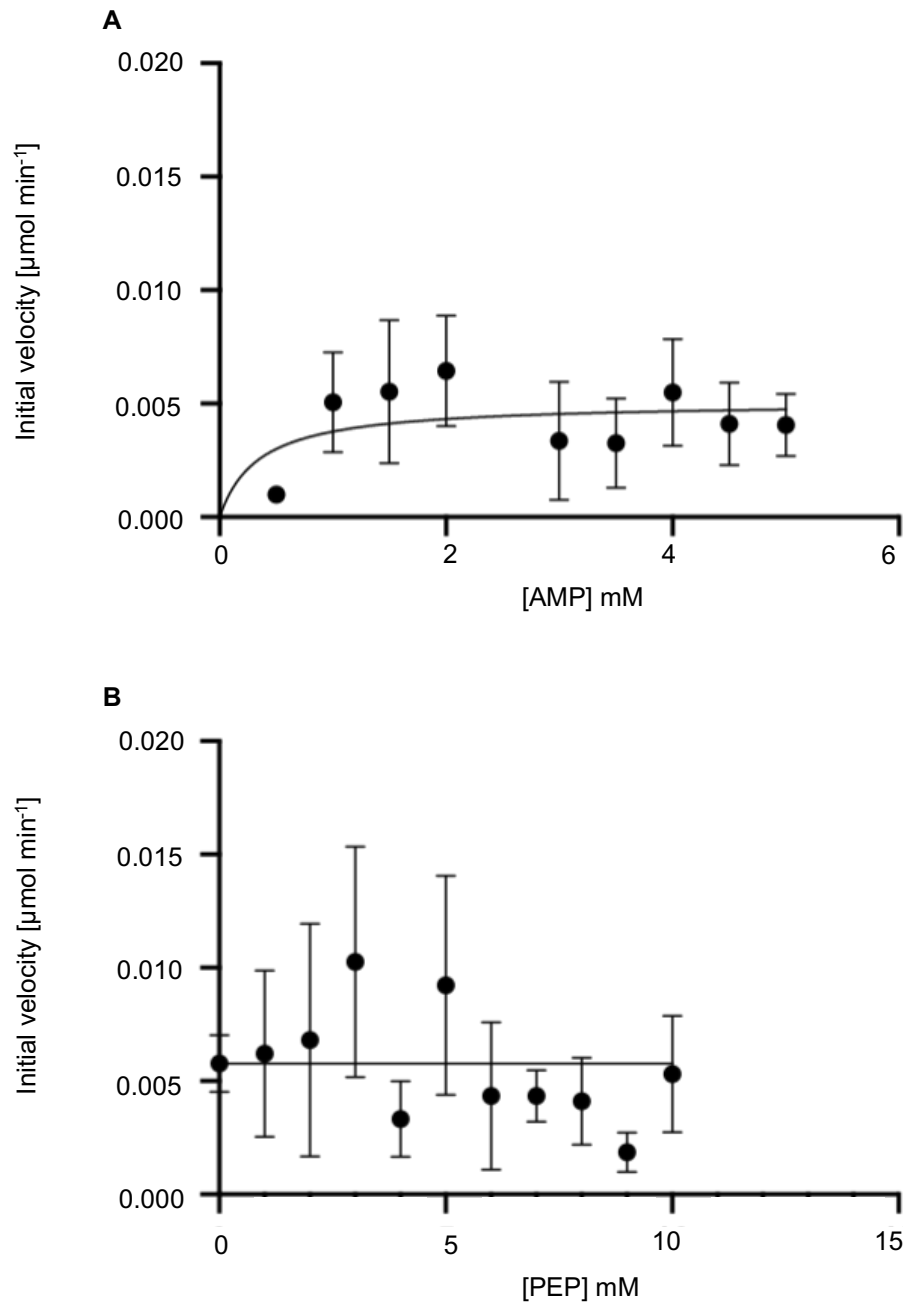


Figure 5.22 Enzyme kinetics experiments for *S. coelicolor* PPDK1 (SCO0208) in the reverse PYR forming direction. A and B show PPDK activity in the reverse (pyruvate forming) direction using ATP and PEP as the variable substrate respectively. When AMP was the rate limiting substrate a concentration of 10 mM PEP was used, and when PEP was the rate limiting substrate a concentration of 0.5 mM AMP was used. These data show no trend.

The forward (PEP forming) reaction for PPDK of *S. clavuligerus* showed increasing velocity as the protein concentration was increased, However, due to time constraints and the availability of the reagents used it was not possible to obtain this data for the two *S. coelicolor* PPDK proteins and perform full Michaelis-Menten kinetics which would determine the V_{max} and K_m s of the enzymes.

It is not possible to fit any of the pyruvate forming reaction data to a Michaelis-Menten curve as there is no trend. This suggests that the PPDK are not catalysing the reverse reaction, however it is possible that the reaction could be missing important cofactors for the reaction. One study involving *M. rosea* also observed little activity in the pyruvate-forming direction [147] which could suggest that organisms that possess pyruvate kinases have no need for PPDK to function in the pyruvate forming direction.

5.11 Discussion

This Chapter has outlined the predicted structures of the Streptomyces PPDK proteins. It has also summarised the parameters by which it is possible to heterologously express the PPDK proteins in *E. coli* and by which it is possible to purify these active proteins using IMAC.

It is established in the literature that metabolic variation in prokaryotes is driven by the acquisition of novel enzymes via horizontal gene transfer, and their subsequent incorporation into the organism's metabolic network. These events appear to have facilitated the proliferation of *ppdk* within actinomycetes generally, and *Streptomyces* spp. in particular. While the multiplicity of *ppdk* genes seems likely to have been the result of horizontal gene transfer, it is unclear if the additional copies are functionally redundant or if they have acquired mutations that confer new activity [167].

Alignments of the amino acid sequences from the *Streptomyces* PPDKs and *Z. mays*, *C. symbiosum*, and *T. brucei* show that both the active and binding sites are conserved.

The evidence that these proteins are conserved was strengthened by the predicted structures of Streptomyces PPDK proteins using AlphaFold2 [160]. The predicted structures of these PPDKs are very similar to those already established for *Z. mays*, *C. symbiosum*, and *T. brucei*. As with the plant PPDKs, each PPDK monomer is constructed of three structural domains: a nucleotide binding domain at the N terminus, the central domain containing the catalytic histidine residue, and the C-terminal PEP/pyruvate binding domain [168]. The main variation in these proteins appears to be the orientation the central β sheet of the protein which contains the catalytic His residue. However, it is unclear if this difference in conformation impacts the function of the protein.

Although it is unclear if this difference in conformation impacts the direction in which the enzyme functions (PEP or pyruvate forming) it is noted that there is a similarity between *S. clavuligerus* PPDK (SCLAV1689) and *S. coelicolor* PPDK1 (SCO0208) and the crystal structure of PPDK from *Z. mays* where the central β sheet is oriented toward the C-terminus of the protein which contains the ATP binding domain. It is documented that while the reactions are reversible, the favoured direction of the reaction differs between organisms. The reaction in maize favours PEP formation, which appears common in plants in general, while the reverse reaction (pyruvate forming) is favoured in some microorganisms [157, 158]. This supports the idea that the two *S. coelicolor* PPDK proteins may have differing functions.

Although the predicted structures obtained using AlphaFold2 provide a good starting point for gathering structural information which is rapid and accessible, as well as providing confidence scores, pLDDT is a measure of how well the prediction would compare to an experimental structure. This is no replacement for more traditional methods for resolving structures. Especially as this prediction method relies on being able to align the target to evolutionarily similar proteins, and it has been established that there has been little research into these proteins other than in plants [169].

Although there was a sequence potentially encoding a second *S. clavuligerus* PPDK protein (SCLAV5441), it was decided that this would be discounted from experiments. This is due to the predicted structure of this protein being shorter than the others and missing the PEP/PYR binding domain meaning that it is almost impossible it would be able to function properly as a PPDK. However it has been reported that only the N-terminus of the protein, containing the Arg111 and His514 residues, is required to produce ATP which could explain why this gene has been maintained in the industrial strains [158]. It has also been stated that in some bacterial species this is the purpose of PPDK, to derive ATP by glycolysis [96].

During expression and purification all three PPDK proteins could be visualised on an SDS-PAGE gel at ~96 kDa providing evidence that they are of equal size to the PPDKs described in the literature. The work shown illustrates that it is possible to heterologously express the PPDK proteins from *Streptomyces* in *E. coli*, and that it is possible to express these proteins at a similar temperature to that at which the strains would be usually grown. This Chapter also shows that the proteins can be purified by step elution using a His-trap column with a concentration of 250 mM of imidazole in the elution buffer. However, it is not possible to purify them by gradient elution with the same concentration. The parameters for expression and purification of these proteins have not been identified before, and differ from the expression and purification of the PPDK protein from *M. rosea*. *M. rosea* was expressed heterologously at 37 °C, expression at this temperature was not possible for the *Streptomyces* PPDKs although *M. rosea* is generally cultured at 45 °C, and that enzyme purified using QAE-Sephadex [147].

The proteins were purified in their active form and using the Abcam PPDK activity assay kit it was possible to show that these enzymes had activity in the PEP forming direction. Little research has been carried out into the activity of PPDKs in bacteria, and the literature reports for plants, where PPDKs have most commonly been studied the activity of the proteins can have an enormous range. Here it is shown that all three of the proteins expressed and purified are functional. However, the activities of the three PPDKs in the forward reaction differ greatly, and considerably lower to the cell lysate activity of *M. rosea* when expressed in *E. coli*. *S. coelicolor* PPDK2 is the most active of the three having an activity of 0.025 U/mg compared with 2.35 U/mg when both have been purified using a column [147]. Unfortunately, due to unavailability of reagents and time constraints it was not possible to perform full kinetic characterisation of these proteins which are experiments that would be required to functionally characterise the enzymes.

Chapter 6: Discussion

6.1 General discussion

The World Health Organisation considers antimicrobial resistance (AMR) one of its top priorities as it is recognised as a significant threat to global human health and to achieving UN sustainability goals [170]. Given the rise in AMR, it is clear there is an urgent need to increase productivity of existing antimicrobials and create rapid increases in titre of new antimicrobials to help combat the global AMR emergency something that is addressed in this thesis. Currently around two-thirds of our clinical antimicrobials are produced as specialised metabolites by the bacterial genus *Streptomyces*, in large-scale submerged fermentations, using strains that have been improved empirically through random mutagenesis over many years.

Humans have been unintentionally domesticating microorganisms for thousands of years [171] and more recently industrial strain improvement programmes have intentionally selected microorganisms for improved performance for industry. *Streptomyces* bacteria are the industrial workhorses for production of numerous pharmaceutically useful metabolites such as antimicrobials, immunosuppressives, anti-cancer agents and anthelmintics [70]. Traditionally, strains of *Streptomyces* that are high producers of bioactive specialised metabolites have been developed via random mutagenesis using chemical or UV mutagens, followed by selection for increased product formation and fermenter behaviour [172]. Whilst this approach to strain improvement is successful at yielding high-producing strains, it is often time consuming and can take years to obtain commercially viable results. This reduces the cumulative profits of antimicrobial drug discovery, meaning that on average novel antibiotics brought to market only achieve profit in year 23 of their patent life which is a recognised disincentive for investing in new product discovery [173]. The time-consuming nature of the classical strain improvement approaches often results in genetic instabilities and diminishing returns in terms of production [174–176].

Therefore, a greater insight into the genomic changes that give rise to increased production levels will help in the design of rational approaches to strain improvement and enable the production of next-generation strains. This could then offer the metabolic flexibility required for transitioning industrial processes to more sustainable feedstocks, increasing productivity, and profitability.

6.1.1 Traditional strain development reduces carbon utilisation profile in bacterial strains.

Strain improvement for industrial antibiotic production is an empirical process - largely the result of poorly understood genetic and biochemical mechanisms that give rise to high producing strains. Recent advances in genomics and genome engineering can be coupled with knowledge of the mutations in high-titre strains and be a powerful route to rational strain design. To achieve this there is a need to understand what mutations have given rise to the high producing strains currently in industry. Gluconeogenic nodes of metabolism can be important in the biosynthesis of a wide range of commercially important specialised metabolites such as polyketides, amphenicols and aminoglycosides [175, 177].

Carbon utilisation in the model *Streptomyces* strain (*S. coelicolor* M145), the wild type *S. clavuligerus* strain DSM738 and then two further industrial strains of *S. clavuligerus* (GSK2 and GSK6) which were both derived from DSM738 through random mutagenesis (see Chapter 3). When analysing the utilisation of 96 different carbon sources using the Biolog it was found, unsurprisingly, that the model *S. coelicolor* has the widest substrate utilisation of the strains used in this thesis. When investigating the industrial lineage of *S. clavuligerus* strains, it was observed that strain DSM738 utilises the widest variety of carbon sources, and this profile narrows with GSK2 (the first industrial strain) and narrows further with GSK6, the most recent

industrial strain used. It is unsurprising that the carbon utilisation narrows like this as the strains have emerged through a random mutagenesis pipeline and are selected for on specific industrial media. However it is important to maintain the ability to use a variety of carbon sources as there is ever growing pressure to use sustainable feedstocks [178, 179]. Similarly, when adapting to a new environment with a single carbon source and single nitrogen source, it has been shown that nutrient utilisation narrows in *E. coli* [137]. Generally in *S. clavuligerus*, as the lineage progresses, there is a reduced ability to utilise certain substrates. The continued utilisation of monosaccharides, some organic acids and increased utilisation of gluconeogenic substrates, amino acids that yield 2-oxoglutarate and oxaloacetate, and oligosaccharides is likely a reflection of the composition of industrial fermentation media.

Previous work had shown that expression of some gluconeogenic genes is upregulated in *S. coelicolor* when medium conditions mimic the kinds of substrates used industrially, with one of the *pyk* genes being upregulated >30-Fold [99]. To investigate the role of these enzymes, a non-standard gluconeogenic route in bacterial mutants was created.

Analysis of the carbon utilisation profiles of the mutant strains shows that the expression of the *ppdk* genes in both *S. coelicolor* and *S. clavuligerus* does influence the rate at which strains grow on these gluconeogenic carbon sources. However, alteration of expression does not completely abolish utilisation of any carbon source, which is important moving forward to not reduce carbon utilisation further.

6.1.2 PPDK performs a unique role in mediating carbon flux in *Streptomyces* spp.

PPDK performs a unique role in mediating the carbon flux in *Streptomyces*. Although there are other gluconeogenic enzymes in the central carbon metabolism of *Streptomyces*, e.g. phosphoenolpyruvate carboxykinase and phosphoenolpyruvate synthase, in *E. coli* the data showed that when the Δpps mutant is challenged to grow gluconeogenically, adding a streptomycete *ppdk* gene does not restore the function [94, 95]. However, on further investigation specifically into the PPDK proteins from the two *Streptomyces* spp, these proteins do not express at 37 °C so in order to conclude that PPDK cannot perform the same function as Pps this experiment would need to be repeated at an appropriate temperature as Chapter 4 shows that these enzymes are expressed at 26°C. There is rather a lack of data regarding activity of the enzymes involved in the PEP-PYR-OXA node in Actinobacteria [94] and specifically many of these studies seem to overlook PPDK, and focus on the role of the glycolytic enzyme pyruvate kinase (pyk). In *S. coelicolor* it has been noted that Pyk is most active under gluconeogenic conditions [94] which contrasts with the findings by Schniete *et al*, that Pyk is more active when grown on glucose supplemented media than media supplemented with a gluconeogenic carbon source, in that case Tween [40]. When *S. coelicolor* is grown on minimal media supplemented with glucose as the sole carbon source and either *ppdk1* or *ppdk2* (SCO0208 or SCO2494) are overexpressed, this resulted in increased growth rate compared to the wild type, likewise knocking down either *ppdk* gene decreased the growth rate compared to the wild type.

Llamas-Ramires *et al* suggest that the activity (gluconeogenic or glycolytic) of PPDK2 (SCO2494), the ortholog of *S. clavuligerus* PPDK (SCLAV1689), is dependent on the

carbon source utilised and the results shown in Chapter 4 support this [94]. When grown on glucose as a sole carbon source, the overexpression of *ppdk* (SCO2494), increases the growth rate of the strain compared to the wild type. Remarkably, we do not see these changes when the strains are grown on a gluconeogenic carbon source such as Tween 40. These data could reflect a similar situation to that in *E. coli*, where growth on glucose as a sole carbon source can interfere with cAMP signalling to CRP (cAMP signalling protein), because some gluconeogenic flux to PEP is required for adenylate cyclase activity [180]. However a caveat is that, in *E. coli*, this is mediated via the glucose PTS (phosphotransferase system), absent in *Streptomyces*. Interestingly, in *E. coli*, growth on glucose as a sole carbon source results in excretion of acetate (as a metabolic overflow product) and the later reassimilation of this as carbon becomes limiting [181]. When *Streptomyces* are grown on glucose as a sole carbon source, pyruvate and 2-oxoglutarate are excreted and then later reassimilated via activity of phosphoenolpyruvate synthase and phosphoenolpyruvate carboxykinase (*ppsA* and *pckA*), the TCA cycle and the glyoxylate shunt [182, 183]. The activity of PPKs could be functionally compensating for the lack of a *ppsA* or the inactivity of the glyoxylate shunt under gluconeogenic conditions [99]. The cross-complementation experiments in *E. coli* (Chapter 4) perhaps do not support this either. In this study, different results for *ppdk1* (SCO0208) from *S. coelicolor* were observed but this could not be compared to the literature where only the expression levels of *ppdk2* were investigated and *ppdk1* (SCO0208) was overlooked. The study by Llamas-Ramires *et al* is not the only one to overlook the presence of two copies of *ppdk* in *S. coelicolor* [94]. Gubbens *et al* [71] investigated the role of glucose kinase on the control of primary metabolic enzymes in *S. coelicolor* and found that *ppdk2* (SCO2494) was the only gene encoding a gluconeogenic enzyme to not show glucose kinase-dependent expression. Gubbens *et al* suggest that there may be a regulatory

role for PPK in *Streptomyces* and the data produced here suggests that there is potential for modulation of central metabolism by these enzymes [6].

Comparing the findings above to the results in this thesis for growth of the *S. coelicolor* CRISPRi knock down strains *ppdk1* and *ppdk2* on minimal media supplemented with Tween 40 as the gluconeogenic carbon source, differing effects are observed for the two *ppdks*. The CRISPRi knockdown of *ppdk2* appears to have no effect on the growth rate of *S. coelicolor* on the gluconeogenic carbon source Tween. However, an effect on growth rate is observed when expression levels of *ppdk1* are altered (knock-down and an additional copy): growth rates are increased from 0.003 h⁻¹ (the specific growth rate of the Wild-Type) to 0.0075 h⁻¹ and 0.009 h⁻¹ respectively. One possible explanation for this is that during growth on gluconeogenic substrates, PPK is functioning bidirectionally to increase flux in the gluconeogenic direction.

Changing the expression of *ppdk* genes in *Streptomyces* alters the growth phenotype on different carbon sources. This could suggest that more carbon flux in the gluconeogenic direction may provide precursors for antibiotic biosynthesis but to investigate this thoroughly, a flux balance analysis would need to be conducted. Interestingly, the presence of glucose in cultures of *S. coelicolor* appears to repress arginine biosynthesis in a Glk-independent manner [6]. This could have important implications for the production of CA in *S. clavuligerus* via gluconeogenic substrates.

6.1.3 PPK proteins in *Streptomyces spp* are canonical with PPKs characterised in other genera

Heterologous production of the *Streptomyces* PPK proteins in *E. coli* enabled partial characterisation of the enzymes. As for other PPKs, SDS polyacrylamide gel electrophoresis showed that the proteins are ~98 kDa in size [147, 184]. The specific

activity of the three purified proteins varied, implying that they likely have a range of kinetic properties, which is observed in other organisms [147, 184]. Although the PPDK activity kit by Abcam provided data from which to calculate their specific activity, it could be that some samples were not folded correctly. Circular dichroism could be used to review if the proteins are folded correctly. This would be a good future experiment to conduct as an incorrectly folded protein may still be able to perform its function, however this would provide inaccurate data [185].

It was not possible to obtain the K_m or V_{max} for any of the three *Streptomyces* PPDK proteins by assaying for the pyruvate forming reaction (glycolytic direction) when coupled with a lactate dehydrogenase assay as little activity was measured and the data did not fit Michaelis-Menten kinetics. Investigating the pyruvate forming activity of the PPDK proteins gave results with very low levels of activity, likely due to *Streptomyces* having two pyruvate kinases increasing the number of other possible routes by which to convert PEP to pyruvate. Therefore it is possible that the PPDKs of *Streptomyces*, like those in *P. shermanii*, *A. xylinum*, *R. rubum*, and C_4 plants such as Maize [158], are primarily responsible for PEP formation via the PEP forming direction of the PPDK from *S. clavuligerus* (SCLAV1698), which does however show activity an increasing velocity as the concentration of protein is increased. There has been little characterisation of these proteins, and one of the only studies involving PPDK from an actinomycete was published over two decades ago. This study showed that in *Microbispora rosea* there is an almost two-fold difference in the affinity for substrate in the PEP-forming direction compared with the pyruvate-forming direction [147].

6.2 Conclusion

The literature on *ppdk* genes in *Streptomyces* and other bacterial strains is rather limited and these enzymes are mostly thought of as important for plants and parasites. This study, using a combination of molecular genetics, bioinformatics, microbial physiology, and biochemistry, has shown that they play an important role in the physiology of *Streptomyces*. This role is also exemplified by selection in industrial strains for increased expression in high-producing strains. We do however know that these genes are widely conserved throughout many Actinomycete genera and the characterisation performed here leads to the conclusion that they are canonical PPDK enzymes in terms of their activity [100].

The evidence gathered in this thesis suggests that when two copies of *ppdk* are present in the genome of an organism, they are likely to have different and specific physiological roles that can be manipulated to increase precursor availability for high value products, and that orthologs are not redundant as they do not perform exactly the same function in the cell.

6.3 Future work

The work described here, as with most scientific investigations, has led to the emergence of more questions than it has answered. To advance the study of PPDK enzymes in *Streptomyces* there are a number of things that require further investigation.

Firstly, optimisation of the clavulanic acid assay in these *S. clavuligerus* strains to identify the exact effects of PPDK on CA production better. The CA assay performed showed that a more sensitive technique is required, so it would be advisable to employ a HPLC assay rather than the plate reader assay used as HPLC is reportedly up to

five times more sensitive, with detection limits of $10 \mu\text{g.L}^{-1}$ - $20 \mu\text{g.L}^{-1}$ and $50 \mu\text{g.L}^{-1}$ – $10,000 \mu\text{g.L}^{-1}$ respectively [186, 187]. Although the range of detection using a HPLC assay is narrow, the lower detection limit would be advantageous in the case of growing the strains on minimal media.

The strains used are highly adapted for very rich industrial media and are a challenge to grow in minimal media, more suited for detailed physiological studies. To address this a more detailed analysis of the metabolism via flux analysis and RNASeq to complement the Biolog data would be desirable.

As previous work to study primary metabolic enzymes in *Streptomyces* has only considered one PPK or overlooked the enzyme completely, it provides an incomplete picture from which to draw conclusions. To understand the role that PPK plays in the anaplerotic node of metabolism alongside the other metabolic enzymes, a flux balance analysis should be performed under a number of conditions. It would be desirable to study the flux of carbon by utilising isotope labelling to identify the exact direction of the carbon flux. This would help to define under what conditions growth or antibiotic production is being increased, as the findings from *S. coelicolor* suggest that there may be independent roles for each isoenzyme.

If this is the case and there are differences in the structural composition of the PPK2 protein (SCO2494) and its orthologs (SCO0208 and SCLAV1689), it would also be valuable to utilise a Helix instrument, developed by Dynamic Biosensors (<https://www.dynamic-biosensors.com/helix/>) to attempt to establish parameters for efficiency of substrate and cofactor binding and understand any conformational/structural changes associated with activity.

Given that the proteins are relatively easy to purify, it would also be desirable to perform crystallisation trials to attempt to characterise the *Streptomyces* PPK proteins structurally as there are no structural data for these proteins from any Actinobacteria. This would also facilitate the evaluation of the accuracy of the

structural predictions made using AlphFold2 in Chapter 5. Crystallisation trials, if successful, would ideally be used to perform x-ray crystallography which is still the gold standard method for structural elucidation of proteins. However, if this was unsuccessful, cryo electron microscopy is another more accessible way to investigate the structure of these proteins. Although AlphaFold2 and other AI prediction methods are becoming increasingly accurate, they are reliant on other structures of similar proteins already established. pLDDT confidence predictions used by these models are also useful when evaluating these predictions, with high confidence scores having a 80-90% accuracy but these models can only be used to predict structures and are not a replacement for cryo EM or X-ray crystallography [188, 189].

Finally, it would be beneficial to perform a full characterisation of the transcriptional activity of these genes across the lifecycle in *Streptomyces* and the transcriptional landscape of the *ppdk* mutant strains to see the effect of disruption on overall cell physiology. This is particularly important as comparative genomics (performed by Dr John Munnoch in the Hoskisson Group) identified no mutations in the *ppdk* gene and it is likely that a regulatory mutation is responsible for the change in expression of *ppdk* in later industrial strains of *S. clavuligerus*. Comparative genomics identified at least seven regulatory proteins with mutational changes in the industrial lineage, with three of these arising in the lineage between GSK5 and GSK6, which is when the upregulation of *ppdk* occurs. None of these regulatory genes have been studied in *Streptomyces*, so it is unclear which may be responsible for the global changes in gene expression.

The work here just scratches the surface in showing the importance of this primary metabolic enzyme in *Streptomyces* and that it is not just important in plants. It also outlines that this overlooked enzyme plays an important role in the utilisation of

different carbon sources and that it can be manipulated to increase the titre of antibiotics. The future work would enable full characterisation of this enzyme enabling it to be further exploited in industrially relevant strains to potentially increase titres of other high value products.

References:

1. **Chen Y, Smanski MJ, Shen B.** Improvement of secondary metabolite production in *Streptomyces* by manipulating pathway regulation. *Applied Microbiology and Biotechnology* 2010;86:19–25.
2. **Barka EA, Vatsa P, Sanchez L, Gaveau-Vaillant N, Jacquard C, et al.** Taxonomy, Physiology, and Natural Products of Actinobacteria. *Microbiology and Molecular Biology Reviews* 2016;80:1–43.
3. **Mark D, DeWald J, Dorador C, Tucker N, Herron P.** Characterisation of anti-pseudomonad activity of hyper-arid *Micromonospora* species. *Access Microbiology* 2020;2: article 293.
4. **Kamjam M, Sivalingam P, Deng Z, Hong K.** Deep Sea Actinomycetes and Their Secondary Metabolites. *Frontiers in Microbiology* 2017;8: article 760.
5. **Traxler MF, Kolter R.** Natural products in soil microbe interactions and evolution. *Natural Product Reports* 2015;32:956–970.
6. **Kumbhar C, Mudliar P, Bhatia L, Kshirsagar A, Watve M.** Widespread predatory abilities in the genus *Streptomyces*. *Archives of Microbiology* 2014;196:235–248.
7. **Zwarycz AS, Whitworth DE.** Myxobacterial Predation: A Standardised Lawn Predation Assay Highlights Strains with Unusually Efficient Predatory Activity. *Microorganisms* 2023;11: article 398.
8. **Genilloud O.** Actinomycetes: still a source of novel antibiotics. *Natural Product Reports* 2017;34:1203–1232.
9. **Nett M, Ikeda H, Moore BS.** Genomic basis for natural product biosynthetic diversity in the actinomycetes. *Natural Product Reports* 2009;26:1362–1384.

10. **Holmes AH, Moore LSP, Sundsfjord A, Steinbakk M, Regmi S, et al.** Understanding the mechanisms and drivers of antimicrobial resistance. *The Lancet* 2016;387:176–187.
11. **Tenover FC.** Mechanisms of Antimicrobial Resistance in Bacteria. *The American Journal of Medicine* 2006;119:S3–S10.
12. **Palzkill T.** Metallo- β -lactamase structure and function. *Annals of the New York Academy of Sciences* 2013;1277:91–104.
13. **Blair JMA, Webber MA, Baylay AJ, Ogbolu DO, Piddock LJV.** Molecular mechanisms of antibiotic resistance. *Nature Reviews Microbiology* 2015;13:42–51.
14. **York A.** A new general mechanism of AMR. *Nature Reviews Microbiology* 2021;19:283–283.
15. **Moo C-L, Yang SK, Yusoff K, Ajat M, Thomas W, et al.** Mechanisms of Antimicrobial Resistance (AMR) and Alternative Approaches to Overcome AMR. *Current Drug Discovery Technologies* 2020;17:430-447.
16. **Blanco P, Hernando-Amado S, Reales-Calderon JA, Corona F, Lira F, et al.** Bacterial Multidrug Efflux Pumps: Much More Than Antibiotic Resistance Determinants. *Microorganisms* 2016;4: article 14.
17. **Bush K, Bradford PA.** β -Lactams and β -Lactamase Inhibitors: An Overview. *Cold Spring Harbor Perspectives in Medicine* 2016;6: article a025247.
18. **Tooke CL, Hinchliffe P, Bragginton EC, Colenso CK, Hirvonen VHA, et al.** β -Lactamases and β -Lactamase Inhibitors in the 21st Century. *Journal of Molecular Biology* 2019;431:3472–3500.
19. **Galleni M, Lamotte-Brasseur J, Raquet X, Dubus A, Monnaie D, et al.** The enigmatic catalytic mechanism of active-site serine β -lactamases. *Biochemical Pharmacology* 1995;49:1171–1178.

20. **Garau G, García-Sáez I, Bebrone C, Anne C, Mercuri P, et al.** Update of the Standard Numbering Scheme for Class B β -Lactamases. *Antimicrobial Agents and Chemotherapy* 2004;48:2347–2349.
21. **Rotondo CM, Wright GD.** Inhibitors of metallo- β -lactamases. *Current Opinion in Microbiology* 2017;39:96–105.
22. **Takasuka TE, Book AJ, Lewin GR, Currie CR, Fox BG.** Aerobic deconstruction of cellulosic biomass by an insect-associated *Streptomyces*. *Scientific Reports* 2013;3: article 1030.
23. **Bontemps C, Toussaint M, Revol PV, Hotel L, Jeanbille M, et al.** (2013). Taxonomic and functional diversity of *Streptomyces* in a forest soil. *FEMS Microbiology Letters* 2013;342:157–167.
24. **Chater KF.** Recent advances in understanding *Streptomyces*. *F1000Research* 2016;5: article 2795.
25. **Manteca Á, Yagüe P.** *Streptomyces* differentiation in liquid cultures as a trigger of secondary metabolism. *Antibiotics* 2018;7: article 41.
26. **Ou X, Zhang B, Zhang L, Dong K, Liu C, et al.** SarA influences the sporulation and secondary metabolism in *Streptomyces coelicolor* M145. *Acta Biochimica et Biophysica Sinica* 2008;40:877–882.
27. **Hoskisson PA, Fernández-Martínez LT.** Regulation of specialised metabolites in Actinobacteria – expanding the paradigms. *Environmental Microbiology Reports* 2018;10:231–238.
28. **de Lima Procópio RE, da Silva IR, Martins MK, de Azevedo JL, de Araújo JM.** Antibiotics produced by *Streptomyces*. *Brazilian Journal of Infectious Diseases* 2012;16:466–471.
29. **Van Wezel GP, McDowall KJ.** The regulation of the secondary metabolism of *Streptomyces*: New links and experimental advances. *Natural Product Reports* 2011;28:1311–1333.

30. **Seipke RF, Kaltenpoth M, Hutchings MI.** *Streptomyces* as symbionts: an emerging and widespread theme? *FEMS Microbiology Reviews* 2012;36:862–876.
31. **Seipke RF, Barke J, Brearley C, Hill L, Yu DW, et al.** A Single *Streptomyces* Symbiont Makes Multiple Antifungals to Support the Fungus Farming Ant *Acromyrmex octospinosus*. *PLOS ONE* 2011;6: article e22028.
32. **Seipke RF, Barke J, Heavens D, Yu DW, Hutchings MI.** Analysis of the bacterial communities associated with two ant–plant symbioses. *MicrobiologyOpen* 2013;2:276–283.
33. **Cafaro MJ, Poulsen M, Little AEF, Price SL, Gerardo NM, et al.** Specificity in the symbiotic association between fungus-growing ants and protective *Pseudonocardia* bacteria. *Proceedings of the Royal Society B: Biological Sciences* 2010;278:1814–1822.
34. **Kaltenpoth M.** Actinobacteria as mutualists: general healthcare for insects? *Trends in Microbiology* 2009;17:529–535.
35. **Selvin J, Shanmughapriya S, Gandhimathi R, Seghal Kiran G, Rajeetha Ravji T, et al.** Optimization and production of novel antimicrobial agents from sponge associated marine actinomycetes *Nocardioopsis dassonvillei* MAD08. *Applied Microbiology and Biotechnology* 2009;83:435–445.
36. **Taylor MW, Hill RT, Piel J, Thacker RW, Hentschel U.** Soaking it up: the complex lives of marine sponges and their microbial associates. *The ISME Journal* 2007;1:187–190.
37. **Kruasuwan W, Lohmaneeratana K, Munnoch JT, Vongsangnak W, Jantrasuriyarat C, et al.** Transcriptome Landscapes of Salt-Susceptible Rice Cultivar IR29 Associated with a Plant Growth Promoting Endophytic *Streptomyces*. *Rice* 2023;16: article 6.

38. **Sousa JAJ, Olivares FL.** Plant growth promotion by streptomycetes: ecophysiology, mechanisms and applications. *Chemical and Biological Technologies in Agriculture* 2016;3: article 24.
39. **Worsley SF, Newitt J, Rassbach J, Batey SFD, Holmes NA, et al.** *Streptomyces* Endophytes Promote Host Health and Enhance Growth across Plant Species. 2020;86:e01053-20.
40. **Schniete JK, Cruz-Morales P, Selem-Mojica N, Fernández-Martínez LT, Hunter IS, et al.** Expanding primary metabolism helps generate the metabolic robustness to facilitate antibiotic biosynthesis in *Streptomyces*. *mBio* 2018;9:e01028-18.
41. **Chevrette MG, Gutiérrez-García K, Selem-Mojica N, Aguilar-Martínez C, Yañez-Olvera A, et al.** Evolutionary dynamics of natural product biosynthesis in bacteria. *Natural Product Reports* 2020;37:556–599.
42. **Kautsar SA, Blin K, Shaw S, Weber T, Medema MH.** BiG-FAM: the biosynthetic gene cluster families database. *Nucleic Acids Research* 2021;49:D490–D497.
43. **Muralikrishnan, B, Retnakumar, RJ, Ramachandran, R, Vinodh, JS, Nijisha, M, et al.** Spatial separation of *Streptomyces* aerial mycelium during fermentation enhances secondary metabolite production. *bioRxiv*. Epub ahead of print 2020. DOI: 10.1101/2020.05.27.118091.
44. **Bibb M.** The regulation of antibiotic production in *Streptomyces coelicolor* A3(2). *Microbiology* 1996;142:1335–1344.
45. **Faddetta T, Renzone G, Vassallo A, Rimini E, Nasillo G, et al.** *Streptomyces coelicolor* vesicles: many molecules to be delivered. *Applied and Environmental Microbiology* 2022;88:e01881-21.
46. **Heul HU van der, Bilyk BL, McDowall KJ, Seipke RF, Wezel GP van.** Regulation of antibiotic production in Actinobacteria: new perspectives from the post-genomic era. *Natural Product Reports* 2018;35:575–604.

47. **Higgins CE, Kastner RE.** *Streptomyces clavuligerus* sp. nov., a β -Lactam Antibiotic Producer. *International Journal of Systematic Bacteriology* 1971;21:326–331.
48. **Song JY, Jeong H, Yu DS, Fischbach MA, Park HS, et al.** Draft genome sequence of *Streptomyces clavuligerus* NRRL 3585, a producer of diverse secondary metabolites. *Journal of Bacteriology* 2010;192:6317–6318.
49. **Ramirez-Malule H, Junne S, Nicolás Cruz-Bournazou M, Neubauer P, Ríos-Estapa R.** *Streptomyces clavuligerus* shows a strong association between TCA cycle intermediate accumulation and clavulanic acid biosynthesis. *Applied Microbiology and Biotechnology* 2018;102:4009–4023.
50. **Ward JM, Hodgson JE.** The biosynthetic genes for clavulanic acid and cephamycin production occur as a ‘super-cluster’ in three *Streptomyces*. *FEMS Microbiology Letters* 1993;110:239–242.
51. **Chandra G, Chater KF.** Evolutionary flux of potentially *bldA*-dependent *Streptomyces* genes containing the rare leucine codon TTA. *Antonie van Leeuwenhoek* 2008;94:111–126.
52. **Huang J, Shi J, Molle V, Sohlberg B, Weaver D, et al.** Cross-regulation among disparate antibiotic biosynthetic pathways of *Streptomyces coelicolor*. *Molecular Microbiology* 2005;58:1276–1287.
53. **Retzlaff L, Distler J.** The regulator of streptomycin gene expression, StrR, of *Streptomyces griseus* is a DNA binding activator protein with multiple recognition sites. *Molecular Microbiology* 1995;18:151–162.
54. **Ryding JJ, Anderson TB, Champness WC.** Regulation of the *Streptomyces coelicolor* calcium-dependent antibiotic by *absA*, encoding a cluster-linked two-component system. *Journal of Bacteriology* 2002;184:794–805.

55. **White J, Bibb M.** *bldA* dependence of undecylprodigiosin production in *Streptomyces coelicolor* A3(2) involves a pathway-specific regulatory cascade. *Journal of Bacteriology* 1997;179:627–633.
56. **Wietzorrek A, Bibb M.** A novel family of proteins that regulates antibiotic production in streptomycetes appears to contain an OmpR-like DNA-binding fold. *Molecular Microbiology* 1997;25:1181–1184.
57. **Sidda JD, Corre C.** Gamma-butyrolactone and furan signaling systems in *Streptomyces*. *Methods in Enzymology* 2012;517:71–87.
58. **Hsiao NH, Söding J, Linke D, Lange C, Hertweck C, et al.** ScbA from *Streptomyces coelicolor* A3(2) has homology to fatty acid synthases and is able to synthesize γ -butyrolactones. *Microbiology* 2007;153:1394–1404.
59. **Horinouchi S, Beppu T.** A-factor as a microbial hormone that controls cellular differentiation and secondary metabolism in *Streptomyces griseus*. *Molecular Microbiology* 1994;12:859–864.
60. **Hutchings MI, Hoskisson PA, Chandra G, Buttner MJ.** Sensing and responding to diverse extracellular signals? Analysis of the sensor kinases and response regulators of *Streptomyces coelicolor* A3(2). *Microbiology* 2004;150:2795–2806.
61. **Maharjan S, Oh TJ, Lee HC, Sohng JK.** Identification and functional characterization of an *afsR* homolog regulatory gene from *Streptomyces venezuelae* ATCC 15439. *Journal of Microbiology and Biotechnology* 2009;19:121–127.
62. **Floriano B, Bibb M.** *afsR* is a pleiotropic but conditionally required regulatory gene for antibiotic production in *Streptomyces coelicolor* A3(2). *Molecular Microbiology* 1996;21:385–396.

63. **Horinouchi S, Kito M, Nishiyama M, Furuya K, Hong S-K, et al.** Primary structure of AfsR, a global regulatory protein for secondary metabolite formation in *Streptomyces coelicolor* A3(2). *Gene* 1990;95:49–56.
64. **Horinouchi S.** AfsR as an integrator of signals that are sensed by multiple serine/threonine kinases in *Streptomyces coelicolor* A3(2). *Journal of Industrial Microbiology and Biotechnology* 2003;30:462–467.
65. **Sohng JK.** Identification and Functional Characterization of an *afsR* Homolog Regulatory Gene from *Streptomyces venezuelae* ATCC 15439. *Journal of Microbiology and Biotechnology* 2009;19:121–128.
66. **Bibb MJ.** Regulation of secondary metabolism in *streptomycetes*. *Current Opinion in Microbiology* 2005;8:208–215.
67. **Romero-Rodríguez A, Rocha D, Ruiz-Villafán B, Guzmán-Trampe S, Maldonado-Carmona N, et al.** Carbon catabolite regulation in *Streptomyces*: new insights and lessons learned. *World Journal of Microbiology and Biotechnology* 2017;33: article 162.
68. **Pereira CS, Santos AJM, Bejerano-Sagie M, Correia PB, Marques JC, et al.** Phosphoenolpyruvate phosphotransferase system regulates detection and processing of the quorum sensing signal autoinducer-2. *Molecular Microbiology* 2012;84:93–104.
69. **Van Wezel GP, König M, Mahr K, Nothaft H, Thomae AW, et al.** A new piece of an old jigsaw: Glucose kinase is activated posttranslationally in a glucose transport-dependent manner in *Streptomyces coelicolor* A3(2). *Journal of Molecular Microbiology and Biotechnology* 2007;12:67–74.
70. **Fernández-Martínez LT, Hoskisson PA.** Expanding, integrating, sensing and responding: the role of primary metabolism in specialised metabolite production. *Current Opinion in Microbiology* 2019;51:16–21.

71. **Gubbens J, Janus M, Florea BI, Overkleeft HS, Van Wezel GP.** Identification of glucose kinase-dependent and -independent pathways for carbon control of primary metabolism, development and antibiotic production in *Streptomyces coelicolor* by quantitative proteomics. *Molecular Microbiology* 2012;86:1490–1507.
72. **Sevcikova B, Kormanec J.** Differential production of two antibiotics of *Streptomyces coelicolor* A3(2), actinorhodin and undecylprodigiosin, upon salt stress conditions. *Archives of Microbiology* 2004;181:384–389.
73. **Ryu YG, Butler MJ, Chater KF, Lee KJ.** Engineering of primary carbohydrate metabolism for increased production of actinorhodin in *Streptomyces coelicolor*. *Applied and Environmental Microbiology* 2006;72:7132–7139.
74. **Khetan A, Malmberg LH, Kyung YS, Sherman DH, Hu WS.** Precursor and cofactor as a check valve for cephamycin biosynthesis in *Streptomyces clavuligerus*. *Biotechnology Progress* 1999;15:1020–1027.
75. **Ruiz B, Chávez A, Forero A, García-Huante Y, Romero A, et al.** Production of microbial secondary metabolites: Regulation by the carbon source. *Critical Reviews in Microbiology* 2010;36:146–167.
76. **Toro L, Pinilla L, Avignone-Rossa C, Ríos-Esteva R.** An enhanced genome-scale metabolic reconstruction of *Streptomyces clavuligerus* identifies novel strain improvement strategies. *Bioprocess and Biosystems Engineering* 2018;41:657–669.
77. **Ser HL, Law JWF, Chaiyakunapruk N, Jacob SA, Palanisamy UD, et al.** Fermentation conditions that affect clavulanic acid production in *Streptomyces clavuligerus*: A systematic review. *Frontiers in Microbiology* 2016;7: article 522.
78. **Paradkar A.** Clavulanic acid production by *Streptomyces clavuligerus*: Biogenesis, regulation and strain improvement. *Journal of Antibiotics* 2013;66:411–420.

79. **Xia H, Zhan X, Mao X-M, Li Y-Q.** The regulatory cascades of antibiotic production in *Streptomyces*. *World Journal of Microbiology and Biotechnology* 2020;36: article 13.
80. **Martín J-F, Liras P.** Engineering of regulatory cascades and networks controlling antibiotic biosynthesis in *Streptomyces*. *Current Opinion in Microbiology* 2010;13:263–273.
81. **Tahlan K, Anders C, Wong A, Mosher RH, Beatty PH, et al.** 5S Clavam Biosynthetic Genes Are Located in Both the Clavam and Paralog Gene Clusters in *Streptomyces clavuligerus*. *Chemistry and Biology* 2007;14:131–142.
82. **Song JY, Jensen SE, Lee KJ.** Clavulanic acid biosynthesis and genetic manipulation for its overproduction. *Applied Microbiology and Biotechnology* 2010;88:659–669.
83. **Cavallieri AP, Baptista AS, Leite CA, Araujo MLG da C.** A case study in flux balance analysis: Lysine, a cephamycin C precursor, can also increase clavulanic acid production. *Biochemical Engineering Journal* 2016;112:42–53.
84. **Li R, Townsend CA.** Rational strain improvement for enhanced clavulanic acid production by genetic engineering of the glycolytic pathway in *Streptomyces clavuligerus*. *Metabolic Engineering* 2006;8:240–252.
85. **Bernal P, Molina-Santiago C, Daddaoua A, Llamas MA.** Antibiotic adjuvants: identification and clinical use. *Microbial Biotechnology* 2013;6:445–449.
86. **Cook MA, Wright GD.** The past, present, and future of antibiotics. *Science Translational Medicine* 2022;14: article eabo7793.
87. **Etebu E, Arikekpar I.** Antibiotics: Classification and mechanisms of action with emphasis on molecular perspectives. *International Journal of Applied Microbiology and Biotechnology Research* 1997;96:535–541.
88. **Donowitz GR, Mandell GL.** Beta-Lactam Antibiotics. *New England Journal of Medicine* 1988;7:419–426.

89. **Röhl F, Rabenhorst J, Zähler H.** Biological properties and mode of action of clavams. *Archives of Microbiology* 1987;147:315–320.
90. **Todd PA, Benfield P.** Amoxicillin/Clavulanic Acid: An Update of its Antibacterial Activity, Pharmacokinetic Properties and Therapeutic Use. *Drugs* 1990;39:261–307.
91. **Butler MS, Henderson IR, Capon RJ, Blaskovich MAT.** Antibiotics in the clinical pipeline as of December 2022. *The Journal of Antibiotics* 2023;76:431–473.
92. **Barona-Gómez F, Chevrette MG, Hoskisson PA.** (2023). On the evolution of natural product biosynthesis. *Advances in Microbial Physiology | Book series* 2023;83:309–349.
93. **Verpoorte R.** Secondary metabolism. In: *Metabolic Engineering of Plant Secondary Metabolism*, 2000;1–29. Dordrecht: Springer Netherlands.
94. **Llamas-Ramírez R, Takahashi-Iñiguez T, Flores ME.** The phosphoenolpyruvate-pyruvate-oxaloacetate node genes and enzymes in *Streptomyces coelicolor* M-145. *International Microbiology* 2020;23:429–439.
95. **McCormick NE, Jakeman DL.** On the mechanism of phosphoenolpyruvate synthetase (PEPs) and its inhibition by sodium fluoride: potential magnesium and aluminum fluoride complexes of phosphoryl transfer. *Biochemistry and Cell Biology* 2015;93:236–240.
96. **Chastain CJ, Failing CJ, Manandhar L, Zimmerman MA, Lakner MM, et al.** Functional evolution of C4 pyruvate, orthophosphate dikinase. *Journal of Experimental Botany* 2011;62:3083–3091.
97. **Koendjibiharie JG, van Kranenburg R, Kengen SWM.** The PEP-pyruvate-oxaloacetate node: variation at the heart of metabolism. *FEMS Microbiology Reviews* 2021;45: article fuaa061.

98. **Taillefer M, Rydzak T, Levin DB, Oresnik IJ, Sparling R.** Reassessment of the transhydrogenase/malate shunt pathway in *Clostridium thermocellum* ATCC 27405 through kinetic characterization of malic enzyme and malate dehydrogenase. *Applied and Environmental Microbiology* 2015;81:2423–2432.
99. **Schniete JK, Reumerman R, Kerr L, Tucker NP, Hunter IS, et al.** Differential transcription of expanded gene families in central carbon metabolism of *Streptomyces coelicolor* A3(2). *Access Microbiology* 2020;2: article 122.
100. **Slamovits CH, Keeling PJ.** Pyruvate-phosphate dikinase of oxymonads and parabasalids and the evolution of pyrophosphate-dependent glycolysis in anaerobic eukaryotes. *Eukaryotic Cell* 2006;5:148–154.
101. **Kirk S, Avignone-Rossa CA, Bushell ME.** Growth limiting substrate affects antibiotic production and associated metabolic fluxes in *Streptomyces clavuligerus*. *Biotechnology Letters* 2000;22:1803–1809.
102. **Bushell ME, Kirk S, Zhao HJ, Avignone-Rossa CA.** Manipulation of the physiology of clavulanic acid biosynthesis with the aid of metabolic flux analysis. *Enzyme and Microbial Technology* 2006;39:149–157.
103. **Saudagar PS, Singhal RS.** A statistical approach using L(25) orthogonal array method to study fermentative production of clavulanic acid by *Streptomyces clavuligerus* MTCC 1142. *Applied Biochemistry and Biotechnology* 2007;136:345–359.
104. **Kohlstedt M, Becker J, Wittmann C.** Metabolic fluxes and beyond-systems biology understanding and engineering of microbial metabolism. *Applied Microbiology and Biotechnology* 2010;88:1065–1075.
105. **Noda-García L, Barona-Gómez F.** Enzyme evolution beyond gene duplication: A model for incorporating horizontal gene transfer. *Mobile Genetic Elements* 2013;3: article e26439.

106. **Noda-García L, Camacho-Zarco AR, Medina-Ruiz S, Gaytán P, Carrillo-Tripp M, et al.** Evolution of Substrate Specificity in a Recipient's Enzyme Following Horizontal Gene Transfer. *Molecular Biology and Evolution* 2013;30:2024–2034.
107. **Sauer U, Eikmanns BJ.** The PEP—pyruvate—oxaloacetate node as the switch point for carbon flux distribution in bacteria. *FEMS Microbiology Reviews* 2005;29:765–794.
108. **Schniete JK, Selem-Mojica N, Birke AS, Cruz-Morales P, Hunter IS, et al.** Actdes – a curated actinobacterial database for evolutionary studies. *Microbial Genomics* 2021;7:1–7.
109. **Howe K, Bateman A, Durbin R.** QuickTree: building huge Neighbour-Joining trees of protein sequences. *Bioinformatics* 2002;18:1546–1547.
110. MUSCLE < Multiple Sequence Alignment < EMBL-EBI. <https://www.ebi.ac.uk/Tools/msa/muscle/> (accessed 24 September 2023).
111. Waterhouse, AM, Procter JB, Martin DMA, Clamp M, Barton GJ. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 2009;25:1189-1191.
<https://academic.oup.com/bioinformatics/article/25/9/1189/203460?login=false> (accessed 26 September 2023).
112. **Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, et al.** Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* 2011;7: article 539.
113. **Paget MS, Chamberlin L, Atrih A, Foster SJ, Buttner MJ.** Evidence that the extracytoplasmic function sigma factor σE is required for normal cell wall structure in *Streptomyces coelicolor* A3(2). *Journal of Bacteriology* 1999;181:204–211.
114. **Taylor RG, Walker DC, McInnes RR.** *E.coli* host strains significantly affect the quality of small scale plasmid DNA preparations used for sequencing. *Nucleic Acids Research* 1993;21:1677–1678.

115. **Kieser T, Bibb MJ, Buttner MJ, Chater KF, Hopwood DA.** *Practical streptomyces genetics: John innes foundation*. Norwich Research Park, Colney.; 2000;44-61
116. **Feeney MA, Newitt JT, Addington E, Algora-Gallardo L, Allan C, et al.** ActinoBase: tools and protocols for researchers working on *Streptomyces* and other filamentous actinobacteria. *Microbial Genomics* 2022;8: article 000824.
117. **Van Dessel, W., Van Mellaert, L., Geukens, N., Lammertyn, E. and Anné, J.,.** Isolation of high quality RNA from *Streptomyces*. *Journal of microbiological methods*, 2004;58:135-137.
118. **Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, et al.** Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular Systems Biology* 2006;2:2006-0008.
119. **Zhang D, Zhang Z, Unver T, Zhang B.** CRISPR/Cas: A powerful tool for gene function study and crop improvement. *Journal of Advanced Research* 2021;29:207–221.
120. **Bradford MM.** A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* 1976;72:248–254.
121. **Hiltner JK, Hunter IS, Hoskisson PA.** Chapter Four - Tailoring Specialized Metabolite Production in *Streptomyces*. In: Sariaslani S, Gadd GM (editors). *Advances in Applied Microbiology*. Academic Press. pp. 237–255.
122. **Khaleeli N, Li R, Townsend CA.** Origin of the β -Lactam Carbons in Clavulanic Acid from an Unusual Thiamine Pyrophosphate-Mediated Reaction. *Journal of the American Chemical Society* 1999;121:9223–9224.
123. **Doi S, Komatsu M, Ikeda H.** Modifications to central carbon metabolism in an engineered *Streptomyces* host to enhance secondary metabolite production. *Journal of Bioscience and Bioengineering* 2020;130:563–570.

124. **Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, et al.** Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 2013;152:1173–1183.
125. **Cui Y, Xu J, Cheng M, Liao X-K, Peng S.** Review of CRISPR/Cas9 sgRNA Design Tools. *Interdisciplinary Sciences* 2018;10:455-465.
126. **Manghwar H, Li B, Ding X, Hussain A, Lindsey K, et al.** CRISPR/Cas Systems in Genome Editing: Methodologies and Tools for sgRNA Design, Off-Target Evaluation, and Strategies to Mitigate Off-Target Effects. *Advanced Science* 2020;7: article 1902312.
127. **Hong H-J, Hutchings MI, Hill LM, Buttner MJ.** The Role of the Novel Fem Protein VanK in Vancomycin Resistance in *Streptomyces coelicolor*. *Journal of Biological Chemistry* 2005;280:13055–13061.
128. **Huang H, Zheng G, Jiang W, Hu H, Lu Y.** One-step high-efficiency CRISPR/Cas9-mediated genome editing in *Streptomyces*. *Acta Biochim Biophys Sin* 2015;47: 231–243.
129. **Ameruoso A, Villegas Kcam MC, Cohen KP, Chappell J.** Activating natural product synthesis using CRISPR interference and activation systems in *Streptomyces*. *Nucleic Acids Research* 2022;50:7751–7760.
130. **Tian T, Kang JW, Kang A, Lee TS.** Redirecting Metabolic Flux via Combinatorial Multiplex CRISPRi-Mediated Repression for Isopentenol Production in *Escherichia coli*. *ACS Synthetic Biology Journal* 2019;8:391–402.
131. **Song CW, Rathnasingh C, Park JM, Kwon M, Song H.** CRISPR-Cas9 mediated engineering of *Bacillus licheniformis* for industrial production of (2R, 3S)-butanediol. *Biotechnology Progress* 2021;37: article e3072.
132. **Schultenkämper K, Brito LF, Wendisch VF.** Impact of CRISPR interference on strain development in biotechnology. *Biotechnology and Applied Biochemistry* 2020;67:7–21.

133. **Larson MH, Gilbert LA, Wang X, Lim WA, Weissman JS, et al.** CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nature Protocols* 2013;8:2180–2196.
134. **Young T, Li Y, Efthimiou G.** (2020). Olive pomace oil can be used as an alternative carbon source for clavulanic acid production by *Streptomyces clavuligerus*. *Waste and Biomass Valorization* 2020;11:3965–3970.
135. **Efthimiou G, Thumser AE, Avignone-Rossa CA.** A novel finding that *Streptomyces clavuligerus* can produce the antibiotic clavulanic acid using olive oil as a sole carbon source. *Journal of Applied Microbiology* 2008;105:2058–2064.
136. **Pérez-Redondo R, Santamarta I, Bovenberg R, Martín JF, Liras P.** The enigmatic lack of glucose utilization in *Streptomyces clavuligerus* is due to inefficient expression of the glucose permease gene. *Microbiology (Reading)* 2010;156:1527–1537.
137. **Cooper V, Lenski R.** The population genetics of ecological specialization in evolving *Escherichia coli* populations. *Nature* 2000;407:736–9.
138. **Shea A, Wolcott M, Daefler S, Rozak DA.** Biolog Phenotype Microarrays. In: Navid A (editor). *Microbial Systems Biology: Methods and Protocols*. Totowa, NJ: Humana Press. pp. 331–373.
139. **Sánchez S, Chávez A, Forero A, García-Huante Y, Romero A, et al.** Carbon source regulation of antibiotic production. *The Journal of Antibiotics* 2010;63:442–459.
140. **Chao YP, Patnaik RA, Roof WD, Young RF, Liao J.** Control of gluconeogenic growth by *pps* and *pck* in *Escherichia coli*. *Journal of Bacteriology* 1993;175:6939–6944.
141. **Medema MH, Alam MT, Heijne WHM, van den Berg MA, Müller U, et al.** Genome-wide gene expression changes in an industrial clavulanic acid

- overproduction strain of *Streptomyces clavuligerus*. *Microbial Biotechnology* 2011;4:300–305.
142. **Medema MH, Cimermancic P, Sali A, Takano E, Fischbach MA.** A systematic computational analysis of biosynthetic gene cluster evolution: lessons for engineering biosynthesis. *PLOS Computational Biology* 2014;10: article e1004016.
 143. **Hiltner JK, Hunter IS, Hoskisson PA.** Tailoring specialized metabolite production in *Streptomyces*. *Advances in Applied Microbiology* 2015;91:237–255.
 144. **Palazzotto E, Tong Y, Lee SY, Weber T.** Synthetic biology and metabolic engineering of actinomycetes for natural product discovery. *Biotechnology Advances* 2019;37: article 107366.
 145. **Świątek MA, Urem M, Tenconi E, Rigali S, Van Wezel GP.** Engineering of N-acetylglucosamine metabolism for improved antibiotic production in *Streptomyces coelicolor* A3(2) and an unsuspected role of NagA in glucosamine metabolism. *Bioengineered* 2012;3:280–285.
 146. **Tjaden B, Plagens A, Dörr C, Siebers B, Hensel R.** Phosphoenolpyruvate synthetase and pyruvate, phosphate dikinase of *Thermoproteus tenax*: key pieces in the puzzle of archaeal carbohydrate metabolism. *Molecular Microbiology* 2006;60:287–298.
 147. **Eisaki N, Tatsumi H, Murakami S, Horiuchi T.** Pyruvate phosphate dikinase from a thermophilic actinomyces *Microbispora rosea* subsp. *aerata*: purification, characterization and molecular cloning of the gene. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* 1999;1431:363–373.
 148. **Baer GR, Schrader LE.** Stabilization of Pyruvate, Pi Dikinase Regulatory Protein in Maize Leaf Extracts. *Plant Physiology* 1985;77:608–611.

149. **McGuire M, Huang K, Kapadia G, Herzberg O, Dunaway-Mariano D.** Location of the Phosphate Binding Site within *Clostridium symbiosum* Pyruvate Phosphate Dikinase,. *Biochemistry* 1998;37:13463–13474.
150. **Ye D, Wei M, McGuire M, Huang K, Kapadia G, et al.** Investigation of the Catalytic Site within the ATP-Grasp Domain of *Clostridium symbiosum* Pyruvate Phosphate Dikinase *. *Journal of Biological Chemistry* 2001;276:37630–37639.
151. **McWilliam H, Li W, Uludag M, Squizzato S, Park YM, et al.** Analysis Tool Web Services from the EMBL-EBI. *Nucleic Acids Research* 2013;41:W597–W600.
152. **Barton GJ.** Sequence alignment for molecular replacement. *Acta crystallographica. Section D, Biological crystallography* 2008;64:25–32.
153. **Bank RPD.** 1D PFV: 2DIK. <https://www.rcsb.org/sequence/2dik> (accessed 12 July 2023).
154. **Bank RPD.** 1D PFV: 1VBG. <https://www.rcsb.org/sequence/1VBG> (accessed 12 July 2023).
155. **Bank RPD.** RCSB PDB.<https://www.rcsb.org/structure/2X0S> (accessed 13 July 2023).
156. **Cosenza LW, Bringaud F, Baltz T, Vellieux FMD.** The 3.0Å Resolution Crystal Structure of Glycosomal Pyruvate Phosphate Dikinase from *Trypanosoma brucei*. *Journal of Molecular Biology* 2002;318:1417–1432.
157. **Nakanishi T, Nakatsu T, Matsuoka M, Sakata K, Kato H.** Crystal Structures of Pyruvate Phosphate Dikinase from Maize Revealed an Alternative Conformation in the Swiveling-Domain Motion,. *Biochemistry* 2005;44:1136–1144.
158. **Lim K, Read RJ, Chen CCH, Tempczyk A, Wei M, et al.** Swiveling Domain Mechanism in Pyruvate Phosphate Dikinase,. *Biochemistry* 2007;46:14845–14853.
159. **Kufareva I, Abagyan R.** Methods of protein structure comparison. *Methods in Molecular Biology* 2012;857:231–257.

160. **Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, et al.** ColabFold: making protein folding accessible to all. *Nature Methods* 2022;19:679–682.
161. Wilson CJ, Choy WY, Karttunen M. IJMS | FAlphaFold2: A Role for Disordered Protein/Region Prediction? *International Journal of Molecular Sciences* 2022;23:4591/
162. **David A, Islam S, Tankhilevich E, Sternberg MJE.** The AlphaFold Database of Protein Structures: A Biologist's Guide. *Journal of Molecular Biology* 2022;434: article 167336.
163. **Tan JS, Ramanan RN, Ling TC, Mustafa S, Ariff AB.** The role of *lac* operon and *lac* repressor in the induction using lactose for the expression of periplasmic human interferon- α 2b by *Escherichia coli*. *Annals of Microbiology* 2012;62:1427–1435.
164. **Nakanishi T, Ohki Y, Oda J, Matsuoka M, Sakata K, et al.** Purification, crystallization and preliminary X-ray diffraction studies on pyruvate phosphate dikinase from maize. *Acta crystallographica. Section D, Biological crystallography* 2004;60:193–194.
165. Evans HJ, **Wood HG.** The mechanism of the pyruvate, phosphate dikinase reaction. *Proceedings of the National Academy of Sciences* 1968;61:1448–1453.
166. **Reeves RE, Menzies RA, Hsu DS.** The Pyruvate-Phosphate Dikinase Reaction. *Journal of Biological Chemistry* 1968;243:5486–5491.
167. **Glasner ME, Truong DP, Morse BC.** How enzyme promiscuity and horizontal gene transfer contribute to metabolic innovation. *The FEBS Journal* 2020;287:1323–1342.
168. **Minges A, Höppner A, Groth G.** Trapped intermediate state of plant pyruvate phosphate dikinase indicates substeps in catalytic swiveling domain mechanism. *Protein Science* 2017;26:1667–1673.

169. **Filgueiras JL, Varela D, Santos J.** Protein structure prediction with energy minimization and deep learning approaches. *Nature Computational Biology* 2023;22:659–670.
170. **World Health Organization.** Monitoring and evaluation of the global action plan on antimicrobial resistance: framework and recommended indicators.
171. **Steensels J, Gallone B, Voordeckers K, Verstrepen KJ.** Domestication of Industrial Microbes. *Current Biology* 2019;29:R381–R393.
172. **Baltz RH.** Genetic manipulation of secondary metabolite biosynthesis for improved production in *Streptomyces* and other actinomycetes. *Journal of Industrial Microbiology and Biotechnology* 2016;43:343–370.
173. Home | AMR Review. <https://amr-review.org/> (accessed 30 August 2024).
174. **Petkovic H, Cullum J, Hranueli D, Hunter IS, Perić-Concha N et al.** Genetics of *Streptomyces rimosus*, the oxytetracycline producer. *Microbiology and Molecular Biology Reviews* 2006;70:704–728.
175. **Cho HS, Jo JC, Shin C-H, Lee N, Choi J-S, et al.** Improved production of clavulanic acid by reverse engineering and overexpression of the regulatory genes in an industrial *Streptomyces clavuligerus* strain. *Journal of Industrial Microbiology and Biotechnology* 2019;46:1205–1215.
176. **Gravius B, Bezmalinović T, Hranueli D, Cullum J.** Genetic instability and strain degeneration in *Streptomyces rimosus*. *Applied and Environmental Microbiology* 1993;59:2220–2228.
177. **Dang L, Liu J, Wang C, Liu H, Wen J.** Enhancement of rapamycin production by metabolic engineering in *Streptomyces hygroscopicus* based on genome-scale metabolic model. *Journal of Industrial Microbiology and Biotechnology* 2017;44:259–270.

178. **Derkx PM, Janzen T, Sørensen KI, Christensen JE, Stuer-Lauridsen B, et al.** The art of strain improvement of industrial lactic acid bacteria without the use of recombinant DNA technology. *Microbial Cell Factories* 2014;13: article S5.
179. **Lee SY, Kim HU.** Systems strategies for developing industrial microbial strains. *Nature Biotechnology* 2015;33:1061–1072.
180. **Kao KC, Tran LM, Liao JC.** A global regulatory role of gluconeogenic genes in *Escherichia coli* revealed by transcriptome network analysis. *Journal of Biological Chemistry* 2005;280:36079–36087.
181. **Kao KC, Yang YL, Boscolo R, Sabatti C, Roychowdhury V et al.** Transcriptome-based determination of multiple transcription regulator activities in *Escherichia coli* by using network component analysis. *Proceedings of the National Academy of Sciences* 2004;101:641–646.
182. **Hobbs G, Obanye AI, Petty J, Mason JC, Barratt E, et al.** An integrated approach to studying regulation of production of the antibiotic methylenomycin by *Streptomyces coelicolor* A3(2). *Journal of Bacteriology* 1992;174:1487–1494.
183. **Karandikar A, Sharples GP, Hobbs G.** Differentiation of *Streptomyces coelicolor* A3(2) under nitrate-limited conditions. *Microbiology (Reading)* 1997;143:3581–3590.
184. **Bringaud F, Baltz D, Baltz T.** Functional and molecular characterization of a glycosomal PPI-dependent enzyme in trypanosomatids: Pyruvate, phosphate dikinase. *Proceedings of the National Academy of Sciences* 1998;95:7963–7968.
185. **Gill, P., Ranjbar, B.** Circular dichroism techniques: biomolecular and nanostructural analyses-a review. *Chemical biology & drug design* 2009;74:101–120.
186. **Ramirez-Malule H, Junne S, López C, Zapata J, Sáez A, et al.** An improved HPLC-DAD method for clavulanic acid quantification in fermentation broths of

Streptomyces clavuligerus. *Journal of Pharmaceutical and Biomedical Analysis* 2016;120:241–247.

187. **Dai X, Xiang S, Li J, Gao Q, Yang K.** Development of a colorimetric assay for rapid quantitative measurement of clavulanic acid in microbial samples. *Science China Life Sciences* 2012;55:158–163.
188. **Miller JE, Agdanowski MP, Dolinsky JL, Sawaya MR, Cascio D, et al.** *AlphaFold* -assisted structure determination of a bacterial protein of unknown function using X-ray and electron crystallography. *Acta crystallographica. Section D, Biological crystallography* 2024;80:270–278.
189. **Terwilliger TC, Liebschner D, Croll TI, Williams CJ, McCoy AJ, et al.** AlphaFold predictions are valuable hypotheses and accelerate but do not replace experimental structure determination. *Nature Methods* 2024;21:110–116.

Appendix 1:

PM-M1 MicroPlate™ - Carbon and Energy Sources

A1 Negative Control	A2 Negative Control	A3 Negative Control	A4 α -Cyclodextrin	A5 Dextrin	A6 Glycogen	A7 Maltitol	A8 Maltotriose	A9 D-Maltose	A10 D-Trehalose	A11 D-Cellobiose	A12 β -Gentiobiose
B1 D-Glucose-6-Phosphate	B2 α -D-Glucose-1-Phosphate	B3 L-Glucose	B4 α -D-Glucose	B5 α -D-Glucose	B6 α -D-Glucose	B7 3-O-Methyl-D-Glucose	B8 α -Methyl-D-Glucoside	B9 β -Methyl-D-Glucoside	B10 D-Salicin	B11 D-Sorbitol	B12 N-Acetyl-D-Glucosamine
C1 D-Glucosaminic Acid	C2 D-Glucuronic Acid	C3 Chondroitin-6-Sulfate	C4 Mannan	C5 D-Mannose	C6 α -Methyl-D-Mannoside	C7 D-Mannitol	C8 N-Acetyl- β -D-Mannosamine	C9 D-Melezitose	C10 Sucrose	C11 Palatinose	C12 D-Turanose
D1 D-Tagatose	D2 L-Sorbose	D3 L-Rhamnose	D4 L-Fucose	D5 D-Fucose	D6 D-Fructose-6-Phosphate	D7 D-Fructose	D8 Stachyose	D9 D-Raffinose	D10 D-Lactitol	D11 Lactulose	D12 α -D-Lactose
E1 Melibionc Acid	E2 D-Melibiose	E3 D-Galactose	E4 α -Methyl-D-Galactoside	E5 β -Methyl-D-Galactoside	E6 N-Acetyl-Neuraminic Acid	E7 Pectin	E8 Sedoheptulosan	E9 Thymidine	E10 Uridine	E11 Adenosine	E12 Inosine
F1 Adonitol	F2 L- Arabinose	F3 D-Arabinose	F4 β -Methyl-D-Xylopyranoside	F5 Xylitol	F6 Myo-Inositol	F7 Meso-Erythritol	F8 Propylene glycol	F9 Ethanolamine	F10 D,L- α -Glycerol-Phosphate	F11 Glycerol	F12 Citric Acid
G1 Tricarballic Acid	G2 D,L-Lactic Acid	G3 Methyl D-lactate	G4 Methyl pyruvate	G5 Pyruvic Acid	G6 α -Keto-Glutaric Acid	G7 Succinamic Acid	G8 Succinic Acid	G9 Mono-Methyl Succinate	G10 L-Malic Acid	G11 D-Malic Acid	G12 Meso-Tartaric Acid
H1 Acetoacetic Acid (a)	H2 γ -Amino-N-Butyric Acid	H3 α -Keto-Butyric Acid	H4 α -Hydroxy-Butyric Acid	H5 D,L- β -Hydroxy-Butyric Acid	H6 γ -Hydroxy-Butyric Acid	H7 Butyric Acid	H8 2,3-Butanediol	H9 3-Hydroxy-2-Butanone	H10 Propionic Acid	H11 Acetic Acid	H12 Hexanoic Acid