# Building a Synthetic Pathway For Nylon precursor Biosynthesis

David Shaun Frederick Jackson

Department of Biochemical Engineering

University College London

Torrington place

London

WC1E 7JE

A thesis submitted for the degree of

Doctor of Philosophy

2017

Declaration:

I, David Shaun Frederick Jackson confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm this has been indicted in the thesis.

# Abstract

Nylon has been one of the most formative textiles in modern times. It's creation in the 1940's proved the linear structure of polymers which led to the ability to design fabrics with robust and predictable properties. This caused a sensation in the hosiery industry and has since moved on to other markets such as the automotive industry, safety equipment and packaging to name a few. With the production being steady year-on-year at around 3.6 megatonnes, a major concern has arisen over fears for the environmental impact of the chemical manufacturing process. Aside from the bulk disposal of salts and solvents, the amount of nitrous oxides produced in large scale manufacture is of great concern in efforts to reduce global warming. These molecules are over 300 times more powerful of a greenhouse gas than carbon dioxide and so any reduction will have a large impact.

There is growing interest in the development of biological routes as a potential to circumvent these limitations of the traditional chemical routes. There are several reasons for this such as the fact that enzymatic routes rely on much lower temperatures, are typically solvent free processes and the emission of hazardous chemicals is avoided. With recent advances in synthetic biology, metabolic engineering offers the possibility of such a process.

This thesis outlines one such biological route to manufacturing a Nylon 6 precursor, 6-aminocaproic acid (6-ACA). By modifying a metabolic pathway from *Acinetobacter calcoaceticus* NCIMB 9871 and adding a Transaminase, a 5 gene operon has been assembled and expressed in *E.coli* BL21(DE3) which is capable of catalytically converting cyclohexanol into the Nylon 6 precursor *in vivo*. Five clonal variants were created and explored to identify the optimal producer. The best producing clone, pQR1553 achieved a titre of 264.48 mg/L (2mM) from 10mM cyclohexanol when grown on 2xTY media and this was confirmed and quantified by GC-MS. This concentration represents one of the highest producing biological routes from an academic research group for 6-ACA in the published literature. However metabolite analysis showed a very clear bottleneck in the conversion of cyclohexanone to ε-caprolactone and attempts to resolve this were made.

# Contents

# List of figures

# List of tables

# Table of Appendix

# Abbreviations

BVMO - Baeyer-Villiger monooxygenase

CHMO - Cyclohexanone, a member of BVMO family

BDO - Butandial

6-ACA - 6-aminocaproic acid (aka 6-aminohexanoate, 6-aminohexanoic acid)

PMP - pyridoxamine phosphate

PLP - pyridoxal phosphate

BSTFA - N, O-Bistrifluoroacetamide

TMCS - Trimethylchlorosilane

CPEC - circular polymerase extension cloning

TAm - Transaminase

6-HDH - 6-hydroxyhexanoate dehydrogenase

# Impact statement

The body of work presented in this thesis could potentially have an impact on several levels. In an academic, and practical sense, the methods used to generate the successful clones i.e. CPEC and PCR based cloning could be used in future by researchers to create multi-gene expression cassettes. To the best of my knowledge, expressing 6 genes using a single plasmid hasn't been shown before in the department and certainly not using the assembly techniques used. This methodology could mean that larger metabolic pathways could be engineered, with minimal effort, to increase the speed of achieving results. Additionally it could lead to a larger interest in producing bulk chemicals from biological routes. The research presented here achieved good titres of 6-aminocaproic acid however with the implementation of design of experiments and further optimisations I believe it is capable of more. Looking at a wider scope of interest there is a necessity to replace petrochemical sources. This is due to a desperate need to reduce the human impact on the environment, conversion of traditional chemical routes into bio-production to eliminate these very harmful manufacturing processes would greatly help in the fight against global warming. There is also scope to utilise this pathway and others to produce alternative products like methylated versions of monomers which are, by traditional chemistry, very difficult to achieve. By optimising these routes, a gap in the plastics market could be capitalised on which could accelerate the development of these plastics into the mass market.

# Acknowlegements

I'd like to sincerely extend my gratitude to my supervisors Professors John Ward and Helen Hailes for all their help and guidance on this project. Although it was a rather unconventional start to the project, it wouldn't have been possible in this time frame had they not been as accommodating as they have been. Their knowledge and expertise seem endless to me. I'd also like to thank Professor Gary Lye for his help during my first year as a doctoral student in the department.

I have been very fortunate in my time nearly five years in the department to have made many friends. Being amongst such a large cohort of students has created a very pleasant environment. I will always cherish my time here and look back with a smile. I'd like to thank the group of people who started the PhD with me: Joe, Ben, Aisha, Pierre, Nick, Dave and Jordan for being legends in my book! I will forever be in their debt for the support they offered when I was at my lowest point. I'd also like to thank the more recent additions to the departments: Tom, Alex, Nuria, Fair,Trish, Henry, Maria. To all these people whom I've shared much banter, you've made my time amazing. The people working in both Ward and Hailes labs have been incredible and so thanks to Maria, Benjy, Damien, Dragana, Yang, Fiona and Jack. There are too many people to thank. I'd also very much like to thank Jess for all she has done and put up with during this PhD. She has been my rock and I don't say it enough how much I appreciate everything she has done for me and the experiences we have shared.

Finally I'd like to thank my Dad, sister Charlene and Auntie Bridget from the bottom of my heart. There are simply no words to describe how much you mean to me and how much of a support you've been.

This degree and thesis is wholly dedicated in memory of my mum, Susan Jackson, who sadly passed away on the 30th of November 2016. Although not academic, she has taught me the most in life. She showed strength beyond her stature and will forever be the bravest person I will ever know. Without her endless love for me and my sister I wouldn't have been able to achieve this.

# Chapter 1: Introduction

# 1 Introduction to Nylon

## 1.1 Nylon; a gamble in a recession

Nylon is arguably the fabric which created an entirely new industry; it proved polymer properties could be both predicted and engineered. This opened the gates to the first viable commercial application of polymers. It was the fabric which revolutionised the hosiery industry in the 1940's and has continued to provide a diverse portfolio of products ever since. Before its conception, a cellulose based material called Rayon was very popular however its properties of a low melting point and fading when exposed to light made it a difficult material. DuPont recognised the need for improvement and it was a scientist named Wallace Carothers who decided to focus on polymers. Initially he proved that polymers have a linear structure by reacting a diacid and a dialcohol however the resultant melting point was still too low. He then focused on polyamides since he believed these would have a higher melting point and eventually settled on Nylon 6,6; a polymer made from hexamethylene diamine and adipic acid, both of which were commonly and cheaply available (Wolfe, 2008). Manufacture was then rapidly scaled up and the total cost from R&D to full scale was $459 million in today's value. Nylon has remained one of the most popular fabrics for 7 decades and production has remained steady at around 3.6 megatonnes per annum (American Chemical Society, 1995), (*World Nylon Fiber Report - Highlights*, 2013).

### 1.1.1 The chemistry behind stocking

The monomer building blocks can vary in size which therefore lead to many types of polymers exhibiting different properties e.g. 6,10; 6,12; 6,4 etc. However by far the two most popular are Nylon 6,6 and Nylon 6 (Antron, 2013). Therefore only these two will be discussed here.

### 1.1.2 Nylon 6,6

Both monomers (adipic acid and hexamethylene diamine) are manufactured from Benzene, an aromatic ring hydrocarbon commonly found in crude oil and coal. Initially adipic acid is produced from Benzene by using the following routes.

1) Benzene is hydrogenated by passing hydrogen gas over liquid Benzene at 450K and 40 atm of pressure with a nickel catalyst. This is then oxidised at 400K and 10atm of pressure with a cobalt salt catalyst to make KA oil (a mixture of cyclohexanone and cyclohexanol). This KA oil is then oxidised in liquid phase with 60% nitric acid using a copper and vanadium catalyst at 350K. This produces copious amounts of the greenhouse gas, Nitrogen Oxide ($N_2O$) (The Essential Chemical Industry, 2013).

2) Benzene is alkylated with propylene using an Aluminium trichloride catalyst (a Friedel Crafts catalyst) to produce Cumene. This is then oxidised to cumyl-hydroperoxide which is then acid cleaved into phenol and acetone. The phenol is then hydrogenated to KA oil (Bellussi & Perego, 2000). The KA oil is then oxidised in liquid phase with 60% nitric acid.

3) Benzene is selectively hydrogenated to cyclohexene and subsequently hydrated to cyclohexanol. This isn't as thermodynamically stable as it is to make cyclohexane and so a large proportion of this is also produced. However, by performing the reaction in an organic phase with a zinc-coated platinum catalyst yields significantly more cyclohexene.

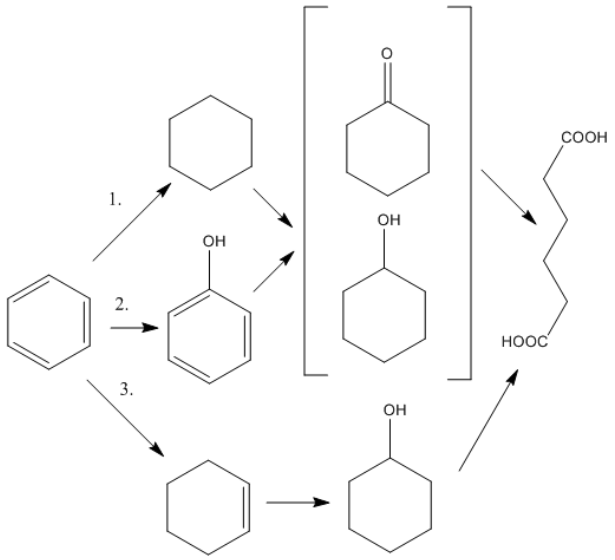



**Figure 1.1 Schematic of the pathways described from benzene to adipic acid. Grouped together in the box is KA oil (Cyclohexanol and cyclohenanone mix). Route 1 benzene is hydrogenated, route 2 benzene is alkylated and route 3 benzene is selectively hydrogenated.**

Adipic acid is also further processed to produce hexamethylene diamine by firstly being converted into adiponitrile by ammoniation and then hydrogenated to hexamethylenediamine (Bellussi & Perego, 2000). Ammoniation is defined as being combined with or treated with ammonia.

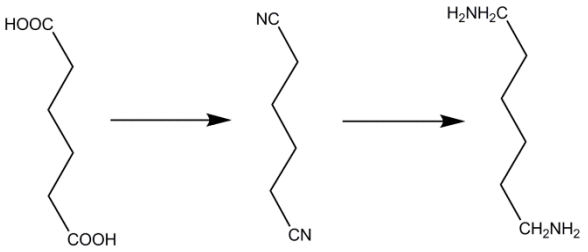



**Figure 1.2 The conversion of adipic acid to hexamethylene diamine. Adipic acid is converted to adiponitrile via ammoniation which is inturn hydrogenated to hexamethylene diamine.**

Finally, after the two monomers are manufactured they are mixed together in equimolar amounts at 550K under pressure in a reaction vessel and they undergo condensation polymerisation.

### 1.1.3 Nylon 6

Developed as a molecule which circumvented the patent DuPont held on Nylon6,6 Nylon 6 (aka polycaprolactam) is a polymer of only one ringed monomer called caprolactam. The process for making caprolactam is similar to adipic acid. It begins with the alkylation of benzene with propylene using a Friedel-crafts catalyst (commonly Aluminium tricloride) to make cumene (or hydrogenated to cyclohexane as in Nylon 6,6 methodology 1 and then oxidised to KA oil). This is then oxidised to cumyl-hydroperoxide and acid cleaved to give phenol and acetone. The phenol is then hydrogenated to KA oil (mixture of cyclohexanone and cyclohexanol) (Bellussi & Perego, 2000). However, pure cyclohexanone is required to make Nylon 6 and so the cyclohexanol must firstly be dehydrated. This is achieved by heating the KA oil to 500K with a copper and chromium oxide catalyst (The Essential Chemical Industry, 2013). The cyclohexanone is then oximated with hydroxylammonium sulphate to cyclohexanone oxime. The final step involves the beckmann rearrangement of cyclohexanone oxime using sulphuric acid into caprolactam. Due to the low pH nature of this reaction, ammonia is added to nutralise this reaction creating a large quiatity of ammonium sulphate bi-product (Bellussi & Perego, 2000).

In order to polymerise caprolactam by ring openning, it is heated to 533K for 12 hours with water, ethanoic acid and nitrogen gas. Due to the instability of caprolactam and the reactivity of its functional groups, the linear molecule reacts with ringed caprolactam and the polymerisation cycle starts again (The Essential Chemical Industry, 2013).

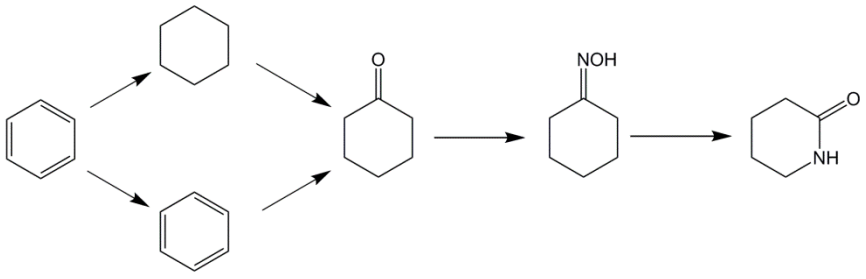



**Figure 1.3 Truncated manufacturing routes to caprolactam from benzene**

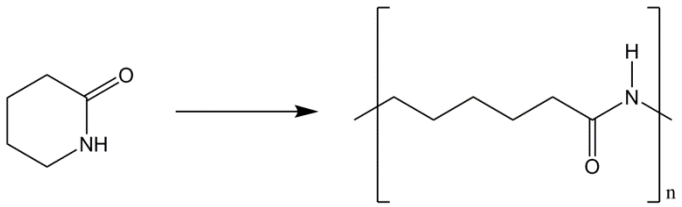



**Figure 1.4 Breaking of the caprolactam ring into a linear monomer is self-catalytic whereby the process generates a linear polymer in one reaction.**

## 1.2 The environmental impact of Nylon production

The traditional industrial chemical synthesis of Nylon requires harsh conditions which, as well as a high cost, are also detrimental to the environment. Firstly, the temperatures required in several steps are very high which places a large energy demand on the process. As the price of energy increases, it directly impacts the cost of goods of Nylon production which which leads to an increasing price for the consumer. Hazardous waste products are also produced in large quantities which is not only costly to dispose of, but are also directly harmful to the environment. For example, low pH steps in the process are often neutralised with ammonia producing ammonium sulphate which is commonly used as a fertiliser. An excess amount of the salt is produced per kilogram of product e.g. 2.8Kg/Kg of cyclohexanone oxime in the ammoximation of cyclohexanol alone (Bellussi and Perego, 2000). However due to its high water solubility, a large proportion of the fertiliser simply runs off from the soil and into lakes and rivers (panap, 2010). This creates a twofold problem. Firstly, water with high nitrogen content nurtures algal growth which quickly stagnates water by metabolising dissolved oxygen and reducing light penetration to aquatic plants. Secondly ammonia, nitrite and nitrate are incredibly toxic to fish and a reduction in bacteria able to metabolise these leads to fatally poisoned fish (panap, 2010).

Arguably the most environmentally detrimental impact of Nylon production is the emission of Nitrous oxides. Nitrous oxides are a powerful greenhouse gas which have an impact 300

times that of $CO_2$ (Ravishankara, Daniel and Portmann, 2009). While the total nitrous production in the chemical industry only contributes 6% to the environmental emissions, owing to its potency, any reduction in production is a great improvement (epa.gov, 2017). In 2002, President Bush issued a policy which committed the US to reducing its greenhouse gas emissions and ever since there has been pressure on other western countries from the government to follow suit (Parker, 2009).

## 1.3 The push towards greener chemistry

During the past decade or so, chemical synthesis companies have begun to face pressure from governments and campaign groups about the effects of waste disposal, safety concerns and reliance on petrochemical feedstocks and their environmental impact (Poliakoff *et al.*, 2002). The traditional chemical synthesis of commodity chemicals and the requirement for the use of finite resources, hazardous solvents, bi-product removal and waste disposal, expensive energy demands for high temperature and pressure steps and the removal of non-desired chiral molecules is placing added pressure on chemical companies. Biocatalysis and synthetic pathways on the contrary have several significant advantages over chemical synthesis. Enzymes, being biologically derived, operate in much milder conditions. They typically require much lower temperatures compared to chemical catalysts and additionally cannot tolerate harsh conditions or solvents. Furthermore many different types can be coupled together in cascades which enable the possibility to use very unstable intermediates because they are simply a substrate for the following enzyme, whereas in chemical synthesis it would prove problematic (Simon *et al.*, 2014). Furthermore, enzymes offer the potential to produce 100% optically pure products therefore, in the case of Nylon 6,6, negating the requirement for chiral separation. It is clear therefore that if biocatalysis is adopted at scale, the potential for large cost savings is

apparent. However, interest from the chemical manufacturing industry has been slow for several reasons. Firstly the industry can be considered quiet conservative and innovative technologies are proving increasingly more difficult to convince company officials to adopt a change. Additionally, the method by which the enzymes are obtained is important since the cost of bulk scale quantities is substantial. Therefore it is probably that non-pure enzymes, for example whole cell catalysts or clarified lysates, are likely to play an important role in producing low-value commodity chemicals. At present benzene is cheap however dwindling oil reserves, political price hikes and reliance on suppliers from unstable countries is driving interest into more innovative, forward thinking technologies that will be required in the next few decades (Bart and Cavallaro, 2015).

## 1.4 Metabolic engineering

Genetic manipulation of microbial cells is a well-established process and something now common place in research. The term ‘’cell factory’’ was coined to describe and imply an engineered cell's ability to manufacture a useful product/function. The emergence of synthetic biology in recent years has in part been aimed at building multi-cascade enzymatic pathways in expression vectors, once a difficult process, is now made with relative ease. The premise of this subsidiary of synthetic biology is to be able to accomplish one of two problems (Bellussi and Perego, 2000); 1) to use relatively cheap feedstocks for engineered microbes and turn these into valuable commodity or fine chemicals or 2) reduce the process economics/the use of hazardous steps required in the chemical synthesis of such chemicals.

This design of such microbes requires knowledge of natural wild-type enzymes, their capabilities and how the product of one enzyme will be the substrate for the next. Whilst natural pathways exist in wild-type bacteria, whether a strain has the ability to produce the

product in industrially relevant quantities is always a concern and hence synthetic biology provides the tools with which to make it so. For example, by introducing an enzyme or a set of enzymes into an expression vector with a powerful promoter, growing the cells at a high density and optimising the growth conditions, a pathway which originates from a slow growing, complex bacterium then potentially operates at a much higher efficiency. Additionally the proteins may themselves be engineered to have an increased affinity for the substrate, to have an increased catalytic rate of reaction or indeed broaden its fidelity in order to make a wider substrate base with which to use as a feed for the pathway. Furthermore, the pathway may be entirely constructed from enzymes taken from several hosts and expressed in the same ''cell factory'' in a combinatorial approach. Whatever the construct, yield is of pertinent importance when thinking about scale-up and industrial relevance.

## 1.5 Efforts towards industrially relevant products

In 2010, a study by Genomatica and the ICIS (International conference on Information systems) showed that 57% of executives in the chemical synthesis industry feel that a move away from petroleum based products is required however the process yields and therefore economics need to match (McGrew, 2010). The momentum in applied metabolic engineering has increased over the last decade which has led to an increase in the number of publications detailing researcher's efforts in achieving high yields of commodity and fine chemicals. Several will be described here. Whilst early research focused on identifying natural producers of chemicals and increasing productivities by random mutagenesis and fermentation optimisation, metabolic engineering is now becoming a useful tool to broaden the scope of bioproduced chemicals (Lee *et al.*, 2011). Akhtar *et al* ,constructed an enzymatic pathway in *E.coli* that conferred the ability to metabolise a broad range of fatty acids into fatty alcohols and alkanes. The pathway consisted of a carboxylic acid reductase,

an aldehyde reductase, a thioesterase and finally a lipase (Akhtar, Turner and Jones, 2013). Reported yields were only 350mg/L, however several parameters needed to be more controlled for instance, the flux through the pathways may need to be adjusted by controlling individual enzyme expression since the $K_{cat}$ of the carboxylic acid reductase is low. Also the enzymes are competing with host variants and so the proportions of fatty alcohol's, alkanes and aldehydes produced varied between batches (Akhtar, Turner and Jones, 2013).

Yim *et al*'s demonstrated outstanding work for the production of 1,4-Butandiol (BDO) from *E.coli.* This compound is currently produced solely from oil and natural gases and is used also to produce polymers. Their approach was to engineer the natural TCA cycle of *E.coli,* and use it together with 4 genes from several different host organisms in order to produce BDO from cheap and renewable carbohydrate feedstocks. By using computer modelling to simulate and predict the metabolic pathway, they narrowed down potential routes based on bi-product formation, yields, pathway length and thermodynamic feasibility. After identifying which enzymes they were going to transform, they then found that in order to balance the pathway flux they needed to arrange the first two genes on a medium copy number plasmid and the last two into a high copy number plasmid. The aim of this being to minimise the burden placed on the cells and so not to disrupt growth characteristics. With this alone they achieved yields of 7.5g/L however they highlighted that bi-product formation was diverting metabolic intermediates away from the pathway. With this in mind they used the software OptKnock to identify target genome encoded host cell enzymes which could be deleted and therefore funnelling into the pathway. They removed 4 genes encoding the following proteins; Alcohol dehydrogenase, pyruvate formate lyse, lactate dehydrogenase and malate dehydrogenase and carbon tracing showed 95% of the glucose was being shunted down the designed pathway. This led to an increase in yield to 18 g/L. They finally recognised that if it was to become commercial, they required a 3-5 fold

increase in yields in order to make economic sense (Yim *et al.*, 2011). However an article by Genomatica in 2010 stated that they have achieved yields of 80g/L BDO production in fermentations ranging from 30-3000L and say that their patented technology matches the economics of traditional chemistry for BDO production (McGrew, 2010).

One hot topic is that of cell engineering research taking place to create a suitable alternative for fuels from cheap and renewable feedstock's such as carbon monoxide, yeast extract and corn steep liquor (Abubackar *et al*. 2015, Abubackar et al. 2012 and Maddipati *et al*. 2011). However in all three examples titres remained low. This could be as a consequence of the emphasis being placed on optimising the fermentation rather than fundamental cell engineering. To substantiate this claim, evidence could be found by Hanai *et al* where they created an *E.coli* K12 strain containing genes for three enzymes cloned from two *Clostridium* species and a *Thermoanaerobacters.* After screening several homologues, the optimum combination was found and yields were reported at 43.5% mol/mol of glucose in shake flask culture and 81.6mM total. This is almost a 3 fold increase on the best known natural producer *C.beijerinckii.* (Hanai, Atsumi and Liao, 2007).

Another example is the microbial production of 1,3-propanediol. This is naturally produced by several bacteria (e.g *Clostridium sp.*) when grown solely on glycerol and in anaerobic conditions to ensure an excess of reduced NADPH is present. 1,-propanediol has been used in a hybrid process with chemically produced terepthalic acid to produce a polymer similar to Nylon. The key enzyme is glycerol dehydratase which relies on a second enzyme, glycerol dehydratase reactivase for vitamin $B_{12}$ reactivation. In this reaction the coenzyme $B_{12}$ is sometimes rendered inactive which ceases catalysis. In this way the GHR as a chaperone to remove inactive $B_{12}$, rendering the glycerol dehydratase capable of biding active vitamin B12. After screening and isolating a library of glycerol dehydratase enzymes, several variants were isolated and cloned which displayed a higher tolerance to substrate, product

and a lower $B_{12}$ dependence (Nakamura and Whited, 2003). To highlight the industrial relevance of 1,3-propanediol, DuPont and Genecor International Inc engineered and patented an *E.coli* K12 strain with a minimal diversion from the designed pathway by several gene deletions and fermentation optimisation. A yield of 130g/L was achieved (Nakamura and Whited, 2003).

Further examples of molecules produced by metabolic engineering include Polylactic acid (Jung and Lee, 2011), the compound GABA which has interests in the food industries and Nylon 4 production (Park *et al.*, 2013), succinic acid (Ho Hong and Yup lee, 2004) and even aldehydes which are not produced naturally in any known living organism since they are converted to their corresponding alcohols rapidly (Kunjapur and Prather, 2015).

## 1.6 The bottlenecks of metabolic engineering

The idea of engineering a host to produce a product it naturally does not produce is rife with problems to overcome. By looking at the literature, several bottlenecks in achieving high, commercially viable yields can be identified. In several of the reports the host's metabolism proved to be a significant blockade by inherently expressing numerous competing enzymes and hence the carbon source metabolism was diverged from the desired pathway. The innate metabolic flux of the host has been fine-tuned by evolution and hence can be a significant obstacle (Na, Kim and Lee, 2010). The effect of such has been shown to be mitigated by either deleting the most probable enzymes and/or coupling with the expression of the synthetic genes from medium/high copy number plasmids. By expressing the proteins in such plasmids it is assumed that the protein concentration out competes natively expressed alternatives as well as the kinetics of the endogenously expressed proteins being such an extent that the conversion rates are smaller than the exogenous proteins. This funnels the majority of the carbon source through the engineered

pathway. Secondly, optimisation of fermentation is an additional requirement. A review by Rosalam and England on Xanthan gum bioproduction detailed the effects of improving culture conditions and identifying optimum feeding strategies not only reduced the fermentation time but also doubled the yield e.g. a feeding strategy of glucose in 5 equal time increments gave a maximum product yield of 18g/L however by enriching the media at 34 hours into the fermentation, yields achieved were 40g/L and this reduced the total time to complete by 20 hours. It is notable that in this particular case the option for strain engineering wasn't possible since any attempt produced lower yields or rendered a lack of growth (Rosalam and England, 2006).

As mentioned above the pathway will be competing with innate enzymes however individual expression levels of each enzyme may be important in yield optimisation. This starts at the ribosome binding site. Variations on the RBS code and distance between it and the start codon greatly affects both the ability of $tRNA^{fmet}$ to bind to the start codon as well as the molecular interaction of the 16s rRNA of the ribosome which therefore impacts on protein expression levels (Na, Kim and Lee, 2010). The fickle nature of the RBS was illustrated by Park *et al* who showed significant protein level increases upon 1 or 2 base changes to the sequence however any modification to the last GAG sequence showed considerably lower expression (Park *et al.*, 2007).

Finally, cofactor availability within a cell may be problematic and limit enzyme activity. The cells enzymes which require $NAD^+$, NADPH, NADH, FAD etc have developed a finely tuned supply chain which minimise wasted energy expenditure. When over-expressing several enzymes in a synthetic pathway, it can be imagined that the supply of the cofactors may not necessarily balance with the demand from the exogenous enzymes.

## 1.7 Nylon precursor research from the literature

Whilst there have previously been notable efforts in chemistry research to produce greener routes to Nylon precursors involving process development and catalytic engineering (Usui & Sato 2003, Thomas & Raja 2005) whole cell biocatalysis and metabolic engineering is rapidly coming of age. Reported titres between the literature are varied and methodologies are broad. Schmidt *et al* developed a pathway in *E.coli* which consisted of an alcohol dehydrogenase followed by a Baeyer-Villegar monoxygenase and then Lipase A to produce poly-ε-caprolactone at 20g/L based on a 200mM feed of cyclohexanol (Schmidt *et al.*, 2015). Their approach was unique in the fact that they expressed ADH from one plasmid in one cell, CHMO from another in another cell, added each type at a ratio of 1:10 respectively and then added the Lipase A into the media. It is notable to mention they used an engineered CHMO for improved stability and reduced product/substrate inhibition capabilities. In addition they reported that isolation of poly-ε-caprolactone involved a simple solvent extraction (Schmidt *et al.*, 2015). In any bioprocess the most costly part of the chain is downstream processing and so by being able to simply extract the product using resins able to bind amine groups is imperative to Nylon precursor biosynthesis scale up.

There have been several attempts to capitalise on the natural tricarboxylic acid cycle by developing pathways to adipic acid bioproduction from glucose. Yu *et al,* developed an *E.coli* strain that co-expressed two plasmids, one containing the gene's *paaj* (β-ketoadipyl-CoA thiolase) and *Ter* (Trans-enoyl-CoA reductase) and another containing the gene's *Hbd* (3-hydroxybuttyrl-CoA dehydrogenase), *Ech* (Enoyl-CoA hydratase), *ptb* (Phosphate butyryltransferase) and *Buk1* (Butyryl kinase). By growing this strain on minimal media supplemented with 10g/L glucose at 30$^{o}$C for 120h and engineering the cell for excess Succinic acid production, titres of 639µg/L were achieved (Yu *et al.*, 2014). An interesting

approach by Frost *et al,* combined both biochemical and chemical methods of manufacture. Firstly they engineered an *E.coli* strain to express two copies of DHS dehydratase from *K.pneumoniae* on the *E.coli* chromosome and a PCA decarboxylase and catechol 1,2-dioxygenase from *A.calcoaceticus* on a plasmid. After 48h of culture a titre of 36.8g/L of *cis, cis*-muconic acid was produced representing a 22% mol/mol yield on D-glucose. Following this the *cis, cis*-muconic acid was treated with activated charcoal and filtered followed by hydrogenation with a 10% platinum on carbon catalyst at 3400 kPa for 2.5 hours achieving a 97% conversion yield to adipic acid (Niu, Draths and Frost, 2002).

1, 5-diaminopentane is naturally produced by *C.glutamicum* and is a precursor to Nylon 5,4 or 5,10. Buschke *et al,* performed a flux analysis on this organism but based on a xylose substrate . They predicted that a maximum yield would result from a strain with high oxidative pentose phosphate pathway activity but zero TCA cycle. After engineering the cell to this design remit, a yield of 100g/L was achieved after 75 hours of culture (Buschke *et al.*, 2013).

Four and five carbon length diamino-alkanes offer an interesting route to molecules which are currently difficult to produce by traditional chemical routes. Kind *et al* developed a strain of the soil bacterium *Corynebacterium glutamicum* which produced Cadaverine (a 5 carbon diamino-alkane) and polymerised it with sebacic acid from castor oil to produce polyamide 5.10. Their approach was three pronged; firstly they cloned a Lysine decarboxylase into a previously engineered high Lysine producing strain, secondly they deleted an N-acetyltransferase which eliminated any side-product formation, and lastly they expressed exporters for lysine and the diaminopentane (Kind *et al.*, 2014). This rather elegant approach meant they achieved yields of 88 g/L with a specific productivity of 2.2 g/L/h. This is perhaps not surprising since their substrate is a natural metabolite of which a large tolerance is inbuilt into the clone. A similar approach was taken when

generating diaminobutane (Aka. putrecine) when Schneider and Wendisch engineered *Corynebacterium glutamicum* with an ornithine decarboxylase and the deletion of a carbamoyltransferase and achieved titres of 6 g/L with a yield of 0.12g/g of glucose (Schneider and Wendisch, 2010).

Clomburg *et al* engineered an *E.coli* strain MG1655 to produce a mixture of hydroxyacids and dicarboxylic acids from glucose or glycerol sources in a multi-enzyme cascade and reported producing adipic acid, 6-hydroxyhexanoic acid, suberic acid, sebacic acid and various dicarboxylic acid chain lengths from $C_6$-$C_{10}$ and reported titres of 0.5 g/L of dicarboxylic acids and 0.8 g/L of hydroxyacids (M.Clomburga *et al.*, 2015).

Hexamethylenediamine (HDMA) has however had significantly less research published to date for bio-based routes. A study by Dros *et al* has suggested that alternative routes towards it's manufacture can exist and outline a route from starch. Their method starts with an enzymatic conversion of starch to high fructose syrup followed by dehydrogenation to 5-(hydroxymethyl)furfural. One of the routes outlines then hydrogenating this to 2,5-Bis(hydroxymethyl)tetrahydrofuran, then hydrogenolysis to 1,6-Hexanediol and finally an amination step to HDMA. Although this was largely an economical and feasibility study, they have stated that this process would cost 20-25% less than the petro-dervied HDMA route as well as a 50% reduction in green house emissions (Dros *et al.*, 2015).

So far the above research discussed in this section have been mostly described using intracellular enzymes in reproducing cells for the reason of co-factor recycling, without any requirement for enzyme purification. However Sattler *et al* has constructed a co-factor self-sufficient pathway *in vitro* producing 6-aminohexanoic acid. This compound can then be treated with superheated steam creating ε-caprolactam. The method they describe uses cyclohexanol as a substrate which is metabolised by an alcohol dehydrogenase, followed by a Baeyer Villiger monoxygenase (BVMO), an esterase, a primary Aldehyde dehydrogenase

(ADH), a ω-Transaminase and finally another esterase. This last step is completed in conjunction with an Alanine dehydrogenase, ammonia and L-alanine. Furthermore they additionally showed that by supplementing with methanol, a cap was placed on side product formation therefore an increase in yield was observed (Sattler *et al.*, 2014). Yields were reportedly 12mM of of 6-aminohexanoic acid however their conversion rates were low which may be resultant of non-optimum buffer conditions or protein instability. Whilst their work clearly is successful in their own right, they encountered protein stability problems from an ADH from *Acinetobacter calcoaceticus* and subsequently used an alternative enzyme. Moreover in terms of scalability, purification and pooling of each enzyme this will not be an ideal process and more than likely will not to be cost effective but useful at lab scale as a proof of principle.

Perhaps the research which would more closely resembles the approach taken in the work presented in this thesis was published in 2015 by Turk *et al*. The group reports cloning six genes on two plasmids which convert 6-alphaketoglutarate from the TCA cycle to 6-ACA. They report titres of 160mg/L of 6-ACA but also state that 2 g/L of total Nylon producing chemicals (i.e. 6-aminovaleric acid, adipic acid, aminopimelate) were achieved in fed-batch lab scale fermentations. They postulated a bottleneck in their pathway which resulted in significant accumulation of a range of bi-products such as 5-aminoadipate and 5 aminovaleric acid and suggested both could be products from α-ketoadipate the result of endogenous aminotransferase for the former biproduct, and a decarboxylase followed by an aminotransferase for the latter (Turk *et al.*, 2015).

So far I have described several research groups and publications in a university setting however this area of metabolic engineering already stretches beyond and into industry. An American company called Verdezyne has developed and patented a yeast cell strain capable of growing on and metabolising all fatty acid compositions to produce adipic acid.

Also being a yeast, it is invulnerable to bacteriophage infections which present a risk to microbial fermentations. They have designed their strain to lack the gene POX4 and amplified genes involved in the β-oxidation pathway and are achieving adipic acid yields of 50g/L in 120h at 300L pilot scale fermentations (Beardslee and Picataggio, 2012). More recently, in November 2014 Verdezyne announced that they had secured a deal with Bio-XCell, a Malaysian biotechnology park focusing on sustainables, to build a facility to manufacture 30 million pounds of dodecandioic acids and adipic acids annually; an unprecedented move in this field (see www.verdezyne.com). However Verdezyne are not alone, there are currently 3 other companies heavily invested in producing sustainable Nylon precursors. Genomatica were recently successful in gaining a patent for producing caprolactone in *E.coli* and hold 7 additional patents. BioAmber is another company producing in yeast however they have licenced their technology and are currently focusing on downstream purifications. Renovia is also another example which has developed *E.coli* strains and is aiming to commercially produce adipic acid by 2018.

## 1.8 *Acinetobacter calcoaceticus* 9871 and the proposed pathway

This study will focus on a synthetic pathway constructed of 4 genes from *Acinetobacter calcoaceticus* coupled together with a Transaminase enzyme that will confer the host with the capacity to metabolise cyclohexanol into 6-aminohexanoic acid (aka 6-aminocaproic acid). A schematic of the pathway is shown in Figure 1.5.

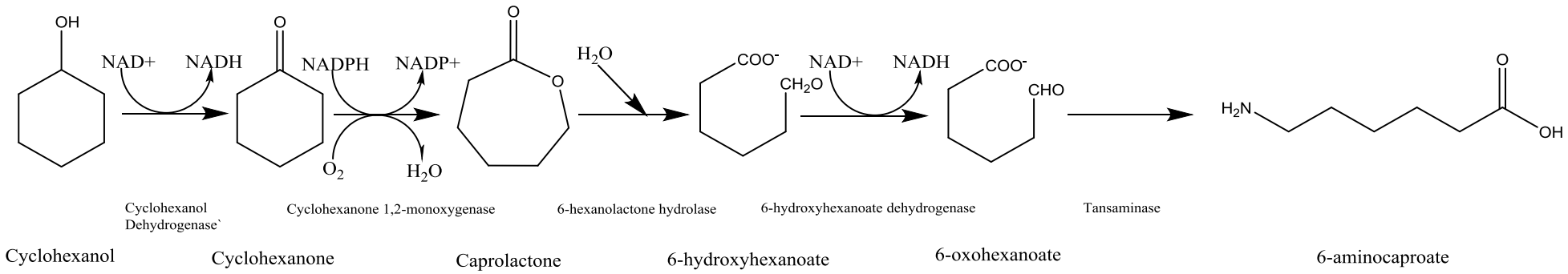



**Figure 1.5 Schematic of the synthetic pathway from cycloexanol to 6-aminocaproate. Cyclohexanol dehydrogenase produces cyclohexanone from cyclohexanol, a cyclohexanone monooxygenase then produces caprolactone (aka BVMO - baeyer villiger monooxygenase), a lactonase (gene - *ChnC*) produces 6-hydroxyhexanoate, a alcohol dehydrogenase (6-hydroxyhexanoate dehydrogenase encoded by *ChnD*) then produces the aldehyde 6-oxohexanoate which is then aminated by a transaminase at the omega position.**

*A.calcoaceticus* was first isolated in 1975 by Donoghue and Trudgill (Donoghue and Trudgill, 1975) and was isolated from soil samples using cyclohexanol as its sole carbon source. Additionally a pathway is proposed which, along with a 5$^{th}$ enzyme (substituting the Transaminase) is capable of producing adipic acid. In the pathway above however that enzyme (*ChnE*) has been removed from the cluster (Figure 1.6) and a Transaminase added to produce 6-aminocaproate. Since this monomer consists of an amine group and a carboxylic group, theoretically the monomers will be able to undergo a dehydration polymerisation and create a Nylon polymer. Genes involved in the same pathway tend to be located close together in clusters and this can be seen in Figure 1.6. It illustrates that *ChnA, ChnC* and *ChnD* are in the reverse orientation to that of *ChnB*. Very little work has been published on this pathway and the only enzyme in the cascade which has been extensively studied is CHMO (a Baeyer-Villager monoxygenase). Research published by Cheng *et al* identified a similar pathway in an *Acinetobacter sp* SE19. Having isolated this strain from an industrial waste water stream and after creating a library of *E.coli* clones harbouring cosmid sections with a length of genomic DNA from *Acinetobacter sp* SE19, they were screened for cyclohexanol oxidation. They identified a gene cluster with high homology to the *Acinetobacter* NCIMB 9871 strain and by random mutagenesis experiments were able to identify gene functions in the 14kb fragment harbouring the pathway (Cheng *et al.*, 2000). Table 1.1 shows a summary of the genes located in that pathway.

| Gene name | Description |
|---|---|
| *ChnA* | Short chain zinc-independent alcohol dehydrogenase. Converts cyclohexanol to cyclohexanone |
| *ChnB* | Authors mutated, caused accumulation of cyclohexanol. Encodes a cyclohexanone monooxygenase (CHMO/BVMO) |
| *ChnC* | Caprolactone hydrolase; converts caprolactone to 6-hydroxyhexanoic acid. |
| *ChnD* | Zinc independent long-chain alcohol dehydrogenase. converts 6-hydroxyhexanoic acid to 6-oxohexanoate. |
| *ChnE* | Aldehyde dehydrogenase; 6-hydroxyhexanoic acid accumulated when mutated. |

**Table 1.1. Summary of the gene's isolated from an *Acinetobacter sp.* by Cheng *et al*.**

Using their gene knock-out variants and analysing the accumulated products they were able to determine a function for the proteins in the pathway. An interesting observation was that when *ChnE* was mutated, instead of 6-oxohexanoate accumulation, 6-hydroxyhexanoic acid accumulated as well as 30% adipic acid (Cheng *et al.*, 2000). Aldehydes are rarely produced intracellularly due to their level of toxicity and so organisms have evolved to be able to rapidly convert these. For example, in *E.coli* BL21 there are 19 aldehyde dehydrogenases (search of the BL21 (DE3) genome at www.kegg.jp/kegg/genome.html). Additionally Cheng *et al* found that the recombinant *E.coli* were unable to grow on cyclohexanol as a sole carbon source (unlike *Acinetobacter calcoaceticus* NCIMB 9871) suggesting other adaptations, perhaps involving the TCA, PPP or β-oxidation pathways which generate ATP and metabolise adipic acid in *Acinetobacter calcoaceticus* NCIMB 9871. They also suggested that aldehydes may react with cellular amines such as amino acids which would explain why none were observed (Cheng *et al.*, 2000).


**Figure 1.6 Schematic to show the gene cluster as it appears in the *Acinetobacter calcoaceticus* 9871 genome. Note that the genes (*ChnA-D*) are not in the same direction but are located in close proximity.**

As mentioned above, *chnB* encodes for a cyclohexanone monooxygenase (CHMO) which is a Baeyer-Villiger monooxygenase (BVMO) a much studied enzyme of great interest in chemistry. Baeyer-Villiger reactions were first discovered over 100 years ago and describe an oxidation of cyclic or acyclic ketones to their corresponding ester or lactone using a peroxide catalyst (Alphand *et al.*, 2003). The oxidant used is often unstable and lacks enantioselectivity therefore has moderate success (Alphand *et al.*, 2003). CHMO is a type 1 flavoenzyme BVMO requiring bound FAD, NADPH and $O_2$ for activity. BVMO's have attracted large interest because of their broad substrate specificity which is capable of producing a range of lactones of great industrial interest. For example, the need for enantioselectivity is never more prevalent than in the synthesis of pharmaceuticals. Monocyclic lactones are small biologically active molecules and bi/polycyclic lactones have shown potential as anti-tumour compounds and cardiac ATPase activators as well as being intermediates in the synthesis of glaucoma and hypertension drugs and even in the manufacture of flavours and fragrances (Alphand *et al.*, 2003). Although there are several well documented papers detailing attempted scale up of these enzymatic reactions, the major hurdle to circumvent is that BVMO's exhibit significant substrate and product inhibition, i.e. is inhibited to a degree by both cyclohexanone and the lactone, albeit a much greater tolerance to lactone (Alphand *et al.*, 2003).

Unfortunately there has been a significant lack of research into the individual gene cluster proteins in the cluster and the only research article published to date looking at the enzyme kinetics is the original article published by Donoghue and Trudgill (Donoghue and Trudgill, 1975). They describe using crude cell extracts and looking for an absorbance change at 340nm in an NADH absorbance assay for *ChnA* oxidation of cyclohexanol to cyclohexanone at pH 10 and then subsequent confirmation by thin layer chromatography. They then subsequently showed that cyclohexanone is rapidly metabolised at pH9 and the only detectable product was 6-hydroxyhexanoic acid. Additionally they suggested that the 3$^{rd}$ enzyme (*ChnC*) had an activity 27 times greater than that of *ChnB* and hence the product of this was rapidly metabolised therefore overcoming product inhibition. Having the knowledge we do now about CHMO's product inhibition, in evolutionary terms it would make sense for *ChnC* to have a high substrate affinity. Finally they identified *ChnD* as a dehydrogenase based on stoichiometric oxygen consumption in the presence of $NAD^+$ and 6-hydroxyhexanoate as a substrate to crude cell lysates. They also found the protein was very labile requiring fresh preparations and even then only exhibiting low activity (Donoghue and Trudgill, 1975).

As can be seen in the schematic of the synthetic pathway (Figure 1.5), the fifth and final enzyme is a ω-Transaminase. Transaminases are a diverse family of enzymes which function to transfer an amine group from a donor to an acceptor (an aldehyde or keto group). This effectively exchanges the Ketone group for the amine group between the two molecules. In nature these enzymes are amongst the most prevalent; every living cell expresses several of these to fulfill roles in interconverting amino acids and their derivatives, to synthesise vitamins and amino sugars for bacterial cell walls etc. They are therefore crucial for nitrogen metabolism (Ward and Wohlgemuth, 2010). Interest in these enzymes is high due to the fact that 70% of marketed pharmaceuticals are made from chiral amino acids. Chemical synthesis is a non-selective method and the absolute

requirement for stereoisomer purity from regulators is helping to drive interest in these enzymes. There are 5 classes of Transaminases in total, all of which use pyridoxal 5'-phosphase (PLP) as a cofactor in the transfer of the amine group to the donor which is then recycled after conversion to Pyridoxamine 5'-phosphate (PMP). Within these classifications, enzymes transfer the amine group to different position's on the acceptor i.e. α Transaminases transfer to position 1, β to position 2, γ to position 3 and ω to the last position. The ω-Transaminases are a class III enzyme and have been subject to significant amount of research in recent times owing largely to their broad substrate specificty. The reactions catalysed by Transaminases are a reversible reaction therefore equilibria are established and can become unfavourable (Ward and Wohlgemuth, 2010). While the characteristics are described in broad terms, some Transaminase varients may exhibit less enzyme activity, less substrate range, lower stability etc and therefore techniques such as random mutagenesis and site-directed mutagenesis have been employed to make improvements (Mathew *et al.*, 2013). Transaminases therefore provide a versatile method for producing a wide variety of compounds and as such may be a popular choice when constructing synthetic pathways.

## 1.9 DNA cloning and operon construction

The emergence of recombinant technology can be considered one of the greatest scientific discoveries of all time and is common place in bioengineering. Through a series of discoveries starting with the observation of DNA transferring between bacteria in the 1920's, followed by the discovery of restriction enzymes and DNA ligases and subsequent plasmid transformation (Pray, 2008) the field of whole cell engineering emerged. Today, various techniques have been developed to allow construction of multi-cistronic plasmids i.e. plasmids expressing several independent proteins. This allows the construction of

complex synthetic pathways on one or more plasmids. A brief description of methods to assemble DNA constructs is below.

### 1.9.1 Restriction digestion and ligation

This is the oldest and most established method of DNA cloning and is restriction cloning. Restrictions enzymes are added to both the DNA and the plasmid vector which recognise and cut specific sites either creating a blunt end or a sticky end with 3' or 5' over hangs. Ideally the digestions should be performed with two different enzymes in the same reaction (aka a double digestion) since this adds directionality (i.e. the gene is inserted in the correct orientation for expression). If the vector is singularly digested it can be dephosphorylated to prevent re-circularisation upon ligation. The DNA segments are then ligated and transformed. This method can be applied to assembling multiple genes in succession however each gene requires its own unique restriction site and it is very laborious requiring many PCR reactions to create the sites at the gene ends, purifications, ligations and transformations.

### 1.9.2 TOPO/Multisite GATEWAY® cloning

Invitrogen's TA/TOPO technology relies on the enzyme DNA Topoisomerase I which has a dual function of restriction enzyme and a ligase. They supply linearised vectors with TOPO covalently bound to the 3' end with an overlapping Thymine residue. The gene to be cloned is then PCR amplified with a *Taq* polymerase which adds a single Adenine to the 3' ends and then both parts are combined. They also have vectors available for directional insertion. Following from this they then developed GATEWAY® which allows rapid transfer of DNA segments by generating overlapping bacteriophage *att* sequences on gene fragments, using proprietary donor vectors which recombine with the gene fragment

creating an entry clone. This entry clone is then recombined with a destination vector and transformed. Initially they designed only 2 *att* sequences however recognising the growing need of researchers to build multi-cistronic pathways they then developed MultiSite GATEWAY® in which they utilise 5 *att* sequences and used the same process to allow construction of up to 4 genes in the same vector (more information is available at www.lifetechnologies.com). There are several drawbacks to this method. Firstly vectors are proprietary from Invitrogen meaning a limitation on expression hosts. Secondly a maximum of only 4 fragments can be assembled together and finally it is very expensive (c.a. £1500/20 reactions).

### 1.9.3 Gibson assembly/In-Fusion HD/HiFI cloning

These three methods are variations of each other and are used to assembly gene constructs together or assemble stretches of DNA. Gene fragments are PCR amplified to contain 15-30 nucleotide flanking regions that are homologous to the DNA stretches which are planned to flank it (see Figure 1.7). The fragments and linearised vector (with overlaps homologous to the 5' gene and the 3' gene) (up to 6 fragments in total) are combined with their master mixes containing a 5' exonuclease which 'chews the DNA back' and creates single strand overlaps. The homology between these overlaps to each other allows them to anneal specifically in the same way PCR primers work. A ligase in the mix then seals the nicks and the assembled plasmid is then transformed. All of this is completed in 15-60 minutes at 50$^{o}$C. All three kits (Gibson from NEB, In-Fusion HD from clontech and HiFi from NEB's own brand) claim to have a high efficiency. Theoretically the fragments should only anneal to the fragments with the homologous overlaps however data published by NEB claim that in a 6 fragment assembly (5 x 1kb fragments + 3.3kb vector), Gibson assembly

achieved a correct arrangement of only 10% of the transformed cells while their HiFi cloning achieved 92% (see www.NEB.com).

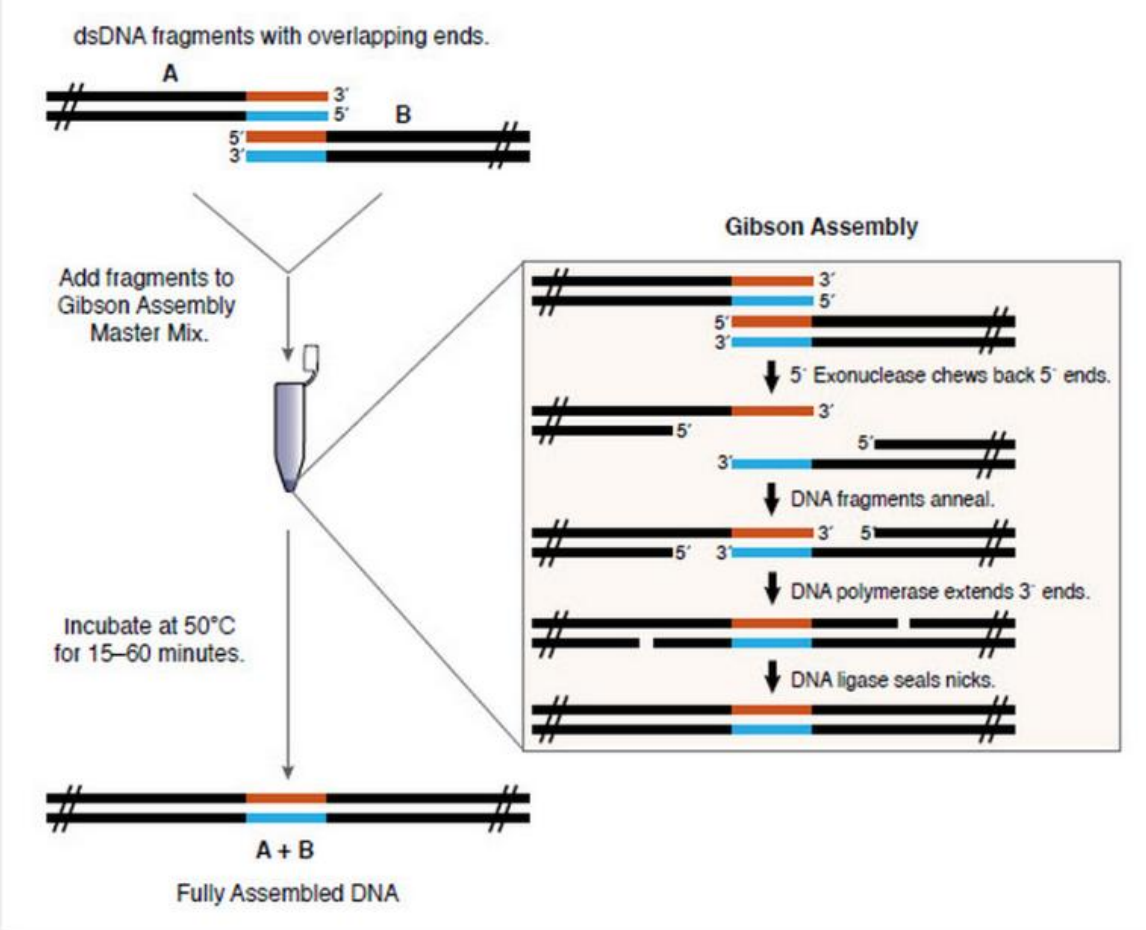
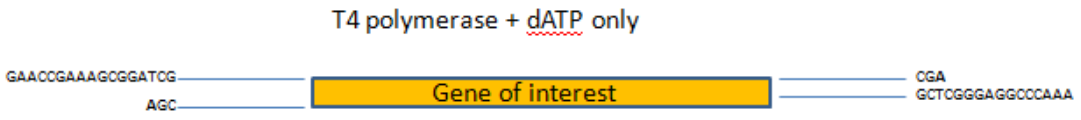



**Figure 1.7 Schematic of the Gibson assembly reaction (image copied from www.NEB.com).**

### 1.9.4 Ligation independent cloning

Ligation independent cloning is another method which uses homologous DNA stretches to anneal sequences in order however the homologous sequences are added to both vector and insert. Firstly using a high fidelity polymerase the overlaps are added. Then a T4 polymerase is used as it exhibits exonuclease activity and after a short time, addition of only one nucleotide is added which halts the exonuclease activity and stalls the polymerase activity at the first occurance of the base which has been added. The fragments are then annealed at 37$^{o}$C and subsequently transformed. An adaptation to this method was developed and named SLIC (sequence and ligation independent cloning). In this method the overlaps can include any nucleotide however RecA (recombinase) must be present. This method was first used by Li and Elledge in 2007 (Li and Elledge, 2007) where they showed that the success of the reaction was dependent upon the length of the homologous region and the number of inserts. They showed that a 5 fragment assembly with 40bp overlaps produced 20/20 correct digestion patterns however two of these had single base mutations obtained from the PCR. In a 10 fragment assembly only 20% were the digested as expected (Li and Elledge, 2007). Although the fidelity of even the worst polymerases is quite good, it is imperative to use a high fidelity polymerase when performing PCR where the amplified fragments will be used in cloning.




**1.8 Schematic of the LIC exonuclease method. After the homologous ovrelaps are added, the exonuclease function of T4 polymerase cuts back until the first base of the only nucleatide present. The polymerase domain takes over and then stalls creating single stranded overhangs.**

### 1.9.5 Circular polymerase extension cloning

This method is the newest and most possibly the most promising method of multi-enzyme cloning available. Published in 2009, Quan and Tian reported the first use of a polymerase based assembly method that requires only overlapping sequences, a thermocycler, a high-fidelity polymerase and dNTP's and showed an impressive success rate (Quan and Tian, 2009). The premise is much the same as Gibson/HiFi/ Phusion by which fragments (including vector) are modified to contain flanking overlapping regions to the segments designed to be either side. The fragments and vector are then combined in ratios of 1-30:1 insert to vector in the presence of a polymerase, dNTP's and buffer and are then cycled (5-30 cycles) through 98$^{o}$C denature, 55$^{o}$C annealing of the primer homologous sequences and finally 78$^{o}$C. They advise that the overlapping segments have as close an annealing temperature as possible (between 60-70$^{o}$C and +/- 2$^{o}$C) with the length of the overlaps of secondary importance (Quan and Tian, 2011). A small volume is then transformed into *E.coli.* Reports suggest the degree of success is dependent upon the individual annealing temperature of the combined overlaps as well as the number of cycles of annealing. CPEC is not an amplification, it simply joins two complimentary fragments and since it is ligase independent, 1 nick per fragment pair is created. This is not a problem however since host DNA repair enzymes seal the nicks *in vivo*(Bryksin and Matsumura, 2010). It is clear that the number of fragments which can be cloned in the same reaction is likely to be limited to 5-10 and is also likely to decrease in efficiency as the number of fragments increases. Following this Bryksin and Matsumura used the same principle but adapted it and showed that a gene can be cloned into a plasmid, anywhere on that plasmid at any specific place, without the need for restriction digestion, purification and ligation. Figure 1.10 shows a schematic of the process. A fragment is again designed with overlaps flanking the left and right of the intended insertion site. The plasmid is then heat denatured, the fragment then anneals anneals and is followed by extension. The next step is to remove any original

template plasmid so the reaction is digested with *DpnI*. The resultant circular plasmid is then transformed (Figure 1.10).

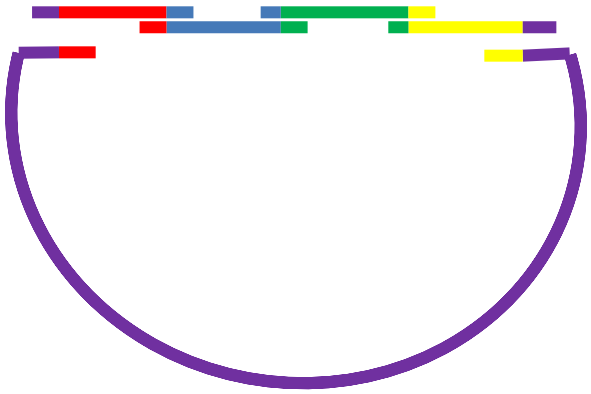


**Figure 1.9 Schematic of CPEC. 4 genes represented by different colours and a purple vector are amplified to contain homologous overlaps to the part it is required to flank.**

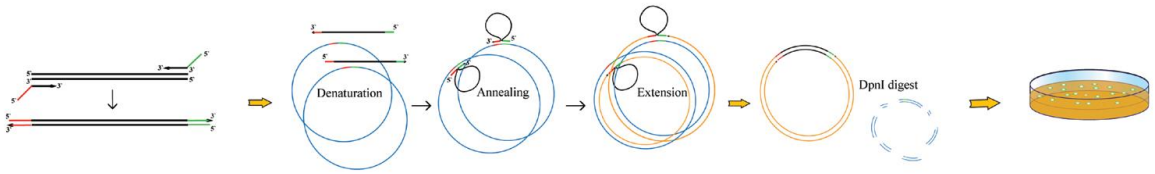



**Figure 1.10 Method developed by Bryksin and Matsumura for modular insertion of a gene anywhere on a plasmid** (image from Bryksin & Matsumura 2010). **DpnI is added to digest the residual template backbone DNA.**

Potential uses of this include extending existing operons or allowing easy deletion and replacement of plasmid genes in a modular fashion, ligation free. They also state that a high concentration of inserts and a low annealing temperature (5-10$^{o}$C below the Tm of the primers) is best for optimal results. Although they achieved colonies with a molar ratio of 5:1 and 50:1 insert to vector, they found that a 250:1 molar ratio between plasmid and insert is optimal (Bryksin and Matsumura, 2010). CPEC therefore offers considerable advantages over other pathway construction methods by being a rapid, versatile and cost effective cloning method.

## 1.10 Codon usage bias

In DNA translation, there are 64 different triplicate sequences coding for only 20 amino acids. This 'degeneracy' has led, through evolution, to different species possessing varying amounts of the individual tRNA's e.g. organism A may prefer to use one codon to code for a particular amino acid yet organism B may prefer to use a different codon with which to recruit the same amino acid (Behura and Severson, 2013). The same can be said for the ability of a ribosome to recognise a ribosome binding site (Park *et al.*, 2007). Part *et al* showed that even slight differences in the Shine-Delgarno sequences and the distance it is from the start site can greatly impact the level of protein expression (Park *et al.*, 2007). Therefore, this must be considered a potential weak link when constructing novel pathways. However, Chen *et al* suggested that the codon usage in *Acinetobacter calcoaceticus* NCIMB 9871 is similar to that of *E.coli* (Chen, Peoples and Walsh, 1988) and may therefore not affect the expression levels of each protein in *E.coli* BL21 (DE3). The basis for their suggestion is that when they were cloning the *Acinetobacter calcoaceticus* NCIMB 9871 BVMO into *E.coli* the expression levels were reasonable.

## 1.11 Project aims

The aim of this project will be to generate create an *E.coli* clone which functionally expresses the *Acinetobacter calcoaceticus* NCIMB 9871 genes coupled with an omega Transaminase obtained from the Ward group's library. This will lead to the *in vivo* conversion of cyclohexanol into 6-aminocaproic acid (6-ACA). The aim will be met by the achievements of several objectives.

- The first objective will be to assemble the four genes (*ChnA, ChnB, ChnC* and *ChnD)* from the *Acinetobacter calcoaceticus* NCIMB 9871 genome into an operon under the

control of a single promoter harbouring an individual ribosome binding sites. This will confer the ability to metabolise cyclohexanol to 6-oxohexanoate.

- The second objective will be to clone each of these individually into *E.coli* BL21 (DE3) cells with a His-tag and assay them for activity indications. The reason for doing this second to assembling the operon is because the operon generation is anticipated to take the most time and will likely be where a large portion of problems will happen.

- Thirdly, a bank of Transaminases, previously cloned by members of our group, will then be screened for optimal activity for conversion of 6-oxohexanoate to the final product 6-aminocaproate (6-aminohexanoate) whilst possessing minimal activity for cyclohexanone. The enzyme showing optimal activity will then be cloned into the pathway.

- The fourth objective will be to develop an analytical method with which to quantify the metabolic intermediates. Given the inherent difficulty analysing amino acids, a robust method will be developed and confirmed by mass spectrometry. Lastly, given enough time, iterations of the plasmid will be attempted in order to minimise the effect of any bottlenecks and feeding strategies will be explored in a simple manner.

# Chapter 2: Materials and Methods

# 2 Materials and methods

## 2.1 General DNA techniques and the 4 *Acinetobacter* sp. gene operon assembly

### 2.1.1 Genome viewer and editor

Genomic sequences were obtained from the NCIB genbank and viewed and edited in either ApE (downloaded from: http://biologylabs.utah.edu/jorgensen/wayned/ape/) or SnapGene viewer (downloaded from: http://www.snapgene.com/products/snapgene_viewer/). Schematics of assembled plasmids were also designed and obtained from these programs.

### 2.1.2 PCR

A nested approach was chosen in which to amplify *chnA, ChnB, ChnC* and *ChnD* from the genome of *Acinetobacter calcoaceticus* NCIMB 9871. Performing PCR on genomic templates can be problematic due to off-target effects which can therefore result in lower yields and many bands on an agarose gel. In this case the 5' and 3' ends of the target genes are low in GC content and the required primers to extract also contain several hairpin loops and 3D structures, hence more appropriate flanking sequences approximately 100bp either side of each gene were chosen. These nested fragments were then used to extract each gene specifically. 20ng of template DNA was used in all PCR reactions. All reagents were sourced from New England Biolabs (USA) and the primers were purchased from Eurofins MWG (Germany). The primers were supplied as a pellet and re-suspended to a concentration of 100mM in DI water. Q5 polymerase was chosen and used in 50 µL reactions. Reactions were set up according to manufactures instructions. Genomic sequences were obtained from Genbank. 2 µL of each PCR reaction was added to 8 µL of water and 2 µL of 6 x Loading dye (NEB, USA) and visualised on agarose gel. The Tm of each gene binding portion of the primer was calculated using Oligo Calc (available at:

http://biotools.nubic.northwestern.edu/OligoCalc.html) and a gradient of +/-5 $^{o}$C was typically chosen. An extension time of 30 s/kb was used on all PCR's at this stage. See appendix for primer sequences used.

### 2.1.3 Agarose gel electrophoresis and gel extraction

DNA samples were separated by agarose gel electrophoreses at 0.8% concentration containing 0.005% (vol/vol) ethidium bromide and running at 120V for 40-50 minutes in 1X TBE buffer (Sigma Aldrich, USA). To confirm whether the reaction had amplified, 2 µL of the PCR reaction was added to 8 µL of water followed by 2 µL of 6X loading dye(NEB, USA). Bands were visualised under UV at 302 nm (Alphaimager mini, Protein simple, USA). If a DNA band was to be purified, the whole PCR reaction mix was loaded with a volume of dye at 1X and ran on a gel. It was excised using a scalpel under UV light and digested using the QiaQuick gel purification kit (Qiagen, Netherlands) following the manufacturer's instructions. To ascertain DNA concentrations, samples were measured for absorbance at 280 nm on a Nanodrop 2000 (Thermoscientific, USA). 10 µL of either 1 Kb or 2- log ladder (NEB, USA) was used as a size reference.

### 2.1.4 Gibson assembly

Each gene fragment concentration was converted into pmol/µL using the equation:

pmol/µL = concentration (ng/µL) x 1000/ (base pairs x 650 daltons)

The protocol described for a 2 (paired inserts A+B, C+D), 4 (4 gene inserts) or 5 fragment assembly (4 inserts $^{+}$ vector) was followed according to manufacturer instructions (NEB) using a total amount of the fragments equivalent to 1 pmol. After addition of the master mix and Di water, the mixture was heated to 50$^{o}$C for 1 hour and then either transformed in 10β cells or visualised on an 0.8% agarose gel.

### 2.1.5 *Restriction digestion and ligation*

For cloning the *Acinetobacter calcoaceticus* NCIMB 9871 genes individually, 1µg of pET29a plasmid (Novagen, Germany) was digested in a 50µL reaction containing 1µL (20units/reaction) of both *NdeI* and *NotI*-HF restriction enzymes (NEB, USA), 5µL of CutSmart buffer (10X; 500mM potassium acetate, 200mM Tris-acetate, 100mM magnesium acetate, 1mg/ml BSA, pH 7.9) for 16 hours. All other digestions were performed with their respective enzymes and digested for 1 hour. Preparative digestions (i.e. to obtain gene fragments for ligation) were carried out in the same 50µL reactions and DNA mass of <1µg. Confirmatory digestions (i.e. post operon construction plasmid miniprep digestions) were carried out in a 10µL volume digesting 200 ng of DNA, 0.2 µL of each restriction enzyme, 1 µL of CutSmart buffer and Di water was added. Digestions were then run on a 0.8% agarose gel with ethidium bromide and visualised under UV. Purifications of the required fragment were done using a QIAquick gel extraction kit (Qiagen, Netherlands). A 3-10 fold molar excess of insert to vector was used in ligation reactions using a Quick ligation$^{Tm}$ Kit (NEB, USA) in a 10 µL reaction following the manufacturers protocol and left to ligate at room temperature for 1 hour and then subsequently either stored at -20$^{o}$C or transformed into a cloning strain.

### 2.1.6 Circular Polymerase Extension Cloning (CPEC)

The methods described by Quan *et al* and Speltz *et al* (Speltz and Regan, 2013) were followed and either equimolar amounts of insert and linearised vector, a 2 fold insert to vector ratio or a 30 fold molar excess to vector approach were tried. A 50 µL reaction was set up containing Q5 polymerase, CutSmart buffer, dNTP's and fragments in the same concentration as for a PCR reaction. The thermocycler procedure of an annealing temperature of 52$^{o}$C for 26 cycles was used in multi-fragment assemblies whilst an annealing temperature of 55$^{o}$C was used at 15 cycles maximum for two fragment assemblies. Following the reaction, samples were either visualised on a 0.8% agarose gel

and the band removed and digested followed by transformation or, 5uL was then transformed directly. Several colonies were grown in 5ml LB broth (Merck Millipore, USA) cultures at 37°C for 16 hours. They were then miniprepped using a QIAprep miniprep kit following the manufacturer's instructions (Qiagen, Hilden, Germany). Plasmids were digested with *HindIII* for 1 hour to confirm the presence of *ChnB* and *ChnD*; *AccI* for *ChnC* ; *PstI* and *NotI* for *ChnA* (NEB, Ipswich, Massachusetts, USA). Correct plasmids were then sequenced by Sanger sequencing (Source Bioscience, Nottingham, UK) for the first and last 800bp (approximately). Successful assemblies were then transformed into BL21 (DE3) (NEB, Ipswich, Massachusetts, USA). The native and non-native RBS operon variants were subsequently named pQR1543 and pQR1548.

## 2.2 General bacterial culture and transformation

### 2.2.1 Plasmid preparation and *E.coli* transformation

From an assembly reaction, 5uL was transformed the cloning strain of choice, *E.coli* 10β cells (NEB, USA) via the heat-shock method following the manufacturer's protocol. A tube of cells was thawed on ice for 10 minutes and the DNA was added to the tube and gently mixed by flicking the tube. This was placed on ice for 30 minutes followed by heating at 42°C for 30 seconds and placing back on ice for 5 minutes. 950 uL of SOC media (supplied with the cells) was added and mixed at 250rpm at 37°C for 1 hour. 200 uL was plated on 50µg/ml Kanamycin LB agar (Merck Millipore, USA) plates and incubated over night at 37°C. If successful, the plate was stored at 4°C for up to 3 weeks. When a plasmid miniprep was required, a single colony was picked and cultured overnight in 5mls of LB media (with 50µg/ml of kanamycin(Sigma-Aldrich, USA)) in a 50ml Falcon tube and incubated at 37°C. LB broth media was made according to manufacturer's recommendations (Merck, Germany) and autoclaved and stored until needed. Following overnight culture, the

suspension was centrifuged at maximum g and at room temperature for 10 minutes to pellet the cells. A QIAprep kit was used (Qiagen, Netherlands) to extract the plasmids following the manufacturers' instructions. 100ng of each plasmid was then transformed via heat shock into the expressing strain *E.coli* BL21 (DE3) (Novagen, Germany) because it encodes an inducible copy of the T7 polymerase required for protein expression from pET29 vectors. pET29a was chosen for cloning because it's a medium copy number plasmid which, when used with BL21 (DE3), has minimal leakiness of protein expression. Transformation of BL21 (DE3) was much the same as 10β cells however they were only heat shocked at 42$^{o}$C for 10 seconds.

### 2.2.2 Routine culture of pQR1553-7

A transformed plate containing colonies stored at 4$^{o}$C for up to 3 weeks. After this time a single colony was picked and re-spread onto a fresh Kanamycin plate, grown at 37$^{o}$C overnight and placed at 4$^{o}$C. For culture experiments to identify pathway metabolites, a single colony was picked and inoculated into 5-10ml of 50ug/ml Kanamycin LB and shaken at 250rpm at 37$^{o}$C overnight. 20-100ml of the chosen media in non-baffled glass flasks was then inoculated to an O.D. of 0.1, grown to 0.6-0.7 and then induced with 1mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) (Sigma-Aldrich, USA).

### 2.2.3 Glycerol stock generation

To make a glycerol stock, single colonies were cultured in 5mls of LB (30µg/ml Kan) overnight. 500µL of the overnight growth was then mixed with 500µL of a 50% glycerol solution (v/v) and stored at -80$^{o}$C and 4.5mls of overnight culture were minipreped and the plasmid stored at -20$^{o}$C.

### 2.2.4 Cell culture growth and pH measurements

Cell cultures were innoculated in 50ml of LB broth in unbaffeled flasks to an O.D of 0.1 and periodically the OD was measured using a spectrophotometer at a wavelength of 600nm.

Dilutions were made to keep the the measurements in the linear range. Cultures were induced at an O.D of 0.6-0.7 and cyclohexanol was added 2 hours later at 10 mM, 30 mM and 50 mM. Culture pH was measured at the same interval using a pH probe.

## 2.3 Protein expression and purification

### 2.3.1 SDS-PAGE

Single colonies of the cell strains expressing either the operon or the individual genes were suspended in 5 ml of LB (30µg/ml kanamycin) overnight and used to inoculate 50mls of LB (30µg/ml Kanamycin) in a 250ml shake flask to an $OD_{600}$ of 0.1 and incubated at 37$^{o}$C and 250 RPM. Once an $OD_{600}$ of 0.6-0.7 has been reached the cells were induced with 50 µL of 1M IPTG at a working concentration of 1mM. 1 ml samples were taken periodically and spun at max g in a bench top centrifuge in microfuge tubes and the supernatant was removed and discarded. The cell pellet was then stored at -20$^{o}$C overnight. Pellets were then thawed and re-suspended in 100 µL of BugBuster (Merck, USA) and incubated at room temperature for 10 mins with gentle shaking to allow adequate time for cell lysis. Samples were then centrifuged at 4$^{o}$C, 12 000g for 20 minutes. 450 µL of 4x Laemmli Sample Buffer (277.8 mM Tris-HCL, pH 6.8, 4.4% LDS, 44.4% (w/v) glycerol, 0.02% bromophenol blue, BIoRad USA) was added to 50 µL of β-mercaptoethanol (Sigma, USA) and mixed at a ratio of 3:1 sample to sample buffer and, together with a broad range 10-250kDa ladder (NEB, USA), boiled at 95$^{o}$C for 5 minutes followed by centrifugation at maximum g for 5 minutes. 10-20 µL's we're loaded onto 12% or 4-20% Mini-PROTEAN®TGX™ (BioRad, USA) gels and run at 180V for 45 minutes. The gel was then washed with RO water (3 x5 minute washes) and gently rocked in 20 mls of InstantBlue protein stain (expedeon, UK) for 1 hour. The gel was washed briefly and then imaged.

### 2.3.2 Protein purification

The cell strain expressing the isolated genes were first plated from a glycerol stocks onto LB agar (30 µg/ml Kan) and grown over night. Single colonies were isolated and grown in 10 ml LB (30µg/ml Kan) over night and this was used to inoculate 100 ml of LB (30 µg/ml Kan) in a 250ml flask to an $OD_{600}$ of 0.1. This was incubated at 37$^{o}$C, 250RPM until an $OD_{600}$ of 0.6-07 and this was then induced with 1mM IPTG and grown for 6 hours. The culture was then removed from the flask and centrifuged at maximum g. The supernatant was discarded and the pellet stored at -20$^{o}$C overnight. The pellet was then re-suspended in BugBuster (Merck, USA) to a volume of 10% the original culture volume and centrifuged at maximum RPM at 4$^{o}$C for 20 minutes. An empty PD-10 column (GE healthcare, UK) was packed with 2 ml of Ni-sepharose 6 Fast flow chromatography medium (GE healthcare, UK) and equilibrated with 10 column volumes of Binding buffer (20mM imidazole, 100mM NaCl, 50mM HEPES, pH 7.5) on ice. The supernatant was firstly passed through a 0.2 µm filter and then loaded onto the column and 2 ml of flow through collected and stored at 4$^{o}$C until needed for SDS-PAGE analysis. The column was then washed with 10 column volumes of binding buffer and 1 ml samples collected. A crude assay to ascertain the presence of protein was performed by diluting Bradfords reagent (Sigma aldrich, USA) 1:4 with water and adding 10 µL of the wash flow through; a blue colour change indicated the presence of protein. Once zero protein was detected, the His-tagged protein was eluted with 10 column volumes of elution buffer (100mM imidazole, 100mM NaCl, 50mM HEPES, pH 7.5). 1 ml fractions were collected in eppendorf tubes (Eppendorf, Germany) and 10 µL samples we're mixed in the crude bradford's assay to test for the presence of eluted protein. Once none could be detected visually, the column was washed with 10 column volumes of RO water and 10 column volumes of 20% Ethanol (vol/vol). It was then stored in this at 4$^{o}$C. The eluted fractions were then kept on ice. The fractions containing the highest concentration of eluted protein we're then pooled to a total volume of 2.5 ml. A PD-10

desalting column (GE healthcare, UK) was used to buffer exchange the protein into 100mM phosphate buffer (pH 7.5) and then the sample was concentrated in a vivaspin 20 (10 kDA cut off, GE healthcare, UK) for 15 minutes in a fixed angle centrifuge or until reduced to approximately 1ml. 10 µL aliquots of all fractions were analysed on an SDS-PAGE to confirm the presence of the target protein. The eluted protein concentration was determined at 280nm using an extinction coefficient calculated from the ExPASy ProtParam algorithm. The elution was then separated into 100 µL aliquots and flash frozen in liquid nitrogen and finally stored at -80$^{o}$C.

## 2.4 *Acinetobacter calcoaceticus* NCIMB 9871 individual enzyme activity assay

To identify enzyme activity, purified protein samples or whole cell lysates were diluted 1:10 and assayed in 100mM Phosphate buffer (pH 7.5) and 0.5 mM of either NADH or NADPH co-factor in a 200 µL reaction. Each substrate was added to a final working concentration of 0.5mM, 5mM, 50mM and 500mM in a 96 well plate. Note that not all enzyme reactions were done at all four of these concentrations. NADH or NADPH absorbance was measured at 340nm in a Tecan plate reader. Cyclohexanol, cyclohexanone and ε-caprolactone were purchased from Sigma-Aldrich (USA) whereas 6-hydroxyhexanoate was purchased from Alfa Aesar (USA).

## 2.5 ω-Transaminase screening and cloning

### 2.5.1 Enzyme generation; deep-well plate culture

28 Ω-Transaminases were screened from a library of recombinant *E.coli* BL21 (DE3) cells expressing TAm's. This library was generated by previous members of the Ward group and stored at -80$^{o}$C for long term storage in glycerol stocks. Fresh glycerol stocks were made in a 96 deep-well plate format. This was done by inoculating 0.8 ml of LB broth (50 µg/ml Kanamycin) in each well of a new PlateOne® deep well plate (Sigma-aldrich, USA) directly

from the glycerol stock using a plastic loop and swabbing the ice. This plate was sealed with a microporous gas permeable Breathe-Easy® sealing membrane (Sigma-aldrich, USA) and grown at 37°C at 250rpm overnight. 0.4 ml of this was then added to 0.4 ml of 50% glycerol in the same pattern as from the original plate. This was then stored at -80°C. To generate reaction plates, a mimic of the glycerol stock plate layout was made by inoculating a new PlateOne® deep well plate (Sigma-aldrich, USA) containing 800µL of 2xTY medium (50µg/ml Kanamycin) per well. The cultures were grown at 37°C, 400 rpm, 85% humidity for 8 h and induced with 1mM IPTG for 16 h at 30°C before harvesting the cells by centrifugation (3000g, 15 min, 4°C). Supernatants were then discarded and the pellets stored at -20°C for up to 1 month or used immediately.

### 2.5.2 6-oxohexanoate synthesis

Two different methods of synthesising 6-oxohexanoate were used. In method 1, cell pellets from 1 ml of induced pChnC and pChnD cultures were lysed using 1x BugBuster (From 10x concentrate; Merck, USA ) in 100mM phosphate buffer with 1 µL/ml Benzonase (Merck, USA). 100 µL of purified caprolactone hydrolase and 6-hydroxyhexanoate dehydrogenase was then added to an assay mix (denoted as assay mix A; 0.625 mM NAD, 1.11 mM pyridoxal 5'-phosphate, 111mM L-Alanine and 45mM ε-caprolactone) in a total volume of 5.5 ml. The reaction was then followed to completion by analysing a 200 µL aliquot at 340 nm on a plate reader (Tecan, UK). Assay mix A was then aliquoted into 180 µL per well into 28 wells on a 96 well plate ready to be used in the ω-TAm reaction. Method 2 involved the acid hydrolysis of methyl-6-oxohexanoate and this was completed by Dr Damien Baud. 80mg of methyl-6-oxohexanoate (Santa Cruz Biotechnology, Inc, Dallas, Tx, USA) was solubilised in 4 mL of $H_2O$/MeOH followed by the addition of 400 mL of an aqueous solution of 10 M NaOH and stirred at room temperature under Argon for 24 hours. The reaction mixture was then acidified using 1M HCL to pH3 and then extracted with three volumes of 15ml dichloromethane. The organic layer was then dried over $MgSO_4$ and

concentrated in a vacuum. The formation of 66mg of yellowish oil was confirmed as 6-oxohexnaoate by NMR. Following this the sample was diluted in methanol to a 50mM stock solution and stored at -20$^{o}$C until required.

### *2.5.3* Screening reaction assay set-up

When using method 1 to synthesise 6-oxohexanoate, as mentioned above, 180 µL of assay mix A was pipetted into each of the 28 wells. When using the chemically synthesised version, 4 µL of the 50 mM stock was added to 176 µL of assay mix B (1.14 mM pyridoxal 5'-phosphate, 114 mM L-Alanine; both purchased from Sigma-Aldrich, USA) in a 96 well plate. Note Assay Mix A and B differ only slightly by concentration since Assay mix A is used at 180 µL and assay mix B is used at 176 µL - both result in the same final concentration of 1 mM and 100 mM of PLP and L-Alanine respectively. Assay mix B also required the addition of NAD since this wasn't required in the TAm reaction with chemically derived 6-oxohexanoic acid. The frozen ω-TAm cell pellet plates were lysed using 50 µL per well of a 1 x BugBuster (Merck, USA) solution in 100 mM phosphate buffer with benzonase and lysed as in section 1.5.2. 20 µL of the TAm lysates was then added to each new reaction mix, sealed and shaken at 750 rpm at 37$^{o}$C for 16 hours. In order to detect potential candidate TAm, when assay mix A was used, 195 µL of the 16 hour reaction was transferred to a new 96 well plate. When assay mix B was used, 185 µL was transferred and added to 10 µL of 10 mM NADH in 100 mM phosphate buffer. In both cases 5 units of Lactate dehydrogenase (Sigma-Aldrich, St. Louis, Missouri, USA) was added to the new plate. The oxidation of NADH (produced from the pChnD reaction in the case of assay mix A) upon catalysis of pyruvate was followed at 340 nm on a plate reader (Tecan, UK).

The ω-TAm screen for activity on cyclohexanone was done using a similar method. 160 µL of assay mix C (1.25 mM PLP, 125 mM L-Alanine in 100 mM phosphate buffer) was added

to 20 µL of 50 mM cyclohexanone (final 5 mM) and 20 µL of ω-TAm cell lysate and incubated at 37°C, 750 rpm for 16 hours. PLP and L-Alanine were again used at 1 mM and 100 mM final concentrations respectively. 175 µL was transferred to a standard 96 well plate. 5 units of LDH and 20 µL of 5 mM NADH was added and read on a plate reader (Tecan, UK) immediately as above.

The methyl-6-oxohexanoate (Santa Cruz Biotechnology, USA) screens were performed by firstly making small aliquotes of 50 mM methyl-6-oxohexanoate in 100 mM phosphate buffer. 4 µL of the 50 mM stock was added to 176 µL of assay mix B (1.14 mM pyridoxal 5'-phosphate, 114 mM L-Alanine; both purchased from Sigma-Aldrich, USA) in a 96 well plate and 20 µL of each ω-TAm cell lysate was used in the 16 hour reaction. 185 µL was transferred to a new standard 96 well plate 10 µL of 10 mM NADH in 100 mM phosphate buffer and 5 units of LDH were added. The reaction was again immediately followed at 340 nm.

### 2.5.4 Cloning in the top three performing ω-Transaminases

pQR810 (containing *Pseudomonas putida* PP_5182) and pQR983 (containing *Deinococcus geothermalis Dgeo*_2743) plasmids were miniprepped from 5 ml overnight cultures as described in 1.2.1. A plasmid list is located in the appendix. Both of these Transaminases were PCR amplified adding overlaps homologous to ChnD and the pet29a vector were added at the 5' and 3' regions respectively. An RBS of AGGAG at position -7 was also added. *BamHI* was used as the 6 bp spacer after the RBS and *NdeI* restriction sites were added immediately after the TAm. See appendix for the list of primers used. The pQR1543 (4 *Acinetobacter calcoaceticus* genes in an operon; natural RBS) and pQR1548 (4 *Acinetobacter calcoaceticus* genes in an operon; artificial RBS) plasmids were inverse PCR linearised (i.e. primers going backwards to linearise the plasmid) and RBS, restriction site

and ω-TAm overlaps were added at flanking ends. An extension time of 30 seconds per kb was used on the TAm genes whilst anything longer than 15 seconds per kb on the inverse PCR would fail. A CPEC reaction was used to assemble the completed pathways using an insert to vector ratio of 5:1 and 10 cycles in the thermocycler annealing at 55$^{o}$C. The reaction mix was directly transformed and colonies were screened by digestion with *BamHI*-HF and *NdeI* (NEB, Ipswich, Massachusetts, USA) as previously described.

*Micrococcus Luteus'* TAm gene MLUT_12025 was amplified from pQR1014 and incorporated the RBS as above. CPEC reactions were tried however all resulted in deletion of the existing genes in the operon. Standard restriction cloning was done on this gene and only into pQR1543. *AvrII* and BbcvI restriction sites were also added at its 5' and 3' regions respectfully due to the presence of the *BamHI* sequence in the gene. The same restriction sites were incorporated into pQR1543 when inverse PCR amplified. The RBS was instead added on to the inverse PCR of pQR1543 at the end of *ChnD* and the *AvrII* site after this. The PCR fragments were purified, ligated and transformed as described earlier.

## 2.6 Assemblying pQR1559; cloning a second T7 promoter, ChnD gene and terminator into pQR1553

Standard restriction cloning was used to add a second copy of *ChnD* under the control of its own T7 promoter, operator and terminator. The operator, T7 promoter and *AvrII* restriction site were added (plus 6 random bases to aid with digestion) to the 5'end of the natural RBS using the nested fragment generated earlier. The reverse primer at this stage included only the 3' end of the gene. The second stage added the terminator to the 3' end of the gene in a single reaction following this. This fragment was then purified and a *SalI* restriction site added to the 3' end of the terminator (again, with 6 random bases added to the 5' end). The pQR1553 plasmid was linearised by inverse PCR at a site between the *LacI*

and *Rop* genes and the *AvrII* site was added at the *Rop* side and *SalI* added at the *LacI* side. This was done using a 15 s/kb extension time as previously described. The two fragments were ligated and cloned as described above. Several miniprepped colonies were digested with *AvrII* and *SalI*-HF® (both from NEB, USA) for 1 hour and bands were observed on agarose gel to confirm the correct assembly. After transforming into *E.coli* BL21 (DE3), the protein profile was examined on SDS-PAGE to identify the presence of all protein bands as well as any possible differences seen with pQR1553. The clone was renamed pQR1559.

## 2.7 Cell culture sample and standard preparation

### 2.7.1 Sample preparations from cultures

Single colonies of the completed pathway clones were inoculated into 5ml of LB supplemented with 50 µg/ml Kanamycin and grown for 16 hours. 50-100ml of fresh LB (or other media) containing 50 µg/ml Kanamycin was then inoculated to an O.D. of 0.1 in non-baffled flasks, grown at 37$^{o}$C shaking at 250 rpm and induced with 1mM IPTG at an O.D of 0.6-0.7. After 2 hours cyclohexanol was fed as a bolus into the flasks which were then left for18 hours. A 1 ml sample was then taken, and sonicated on ice for 2 cycles at 10 µm amplitude for 25 seconds each cycle. The sample was split into two 500 µL aliquots. One was mixed in a 1:1 ratio with Ethyl Acetate, vortexed and mixed thoroughly and then centrifuged at max g for 10 minutes. The top organic layer was then taken and pipetted into a 300 µL chromacol glass vials with inserts and crimped shut (Thermo Fisher Scientific, Waltham, MA, USA). They were then ran on GC-MS immediately (see next section). In order to identify molecules after ε-caprolatone, the second 500 µL sonicated sample was then dried in one of two ways. Initially the samples were left to dry at 75$^{o}$C, ependorf cap open overnight until a dried paste was left. However, results indicated this was suboptimal and

so subsequently smaller samples of 100 µL and concentrated to dryness under vacuum at room temperature using a Jouan RC 10-22, (Thermo Scientific, USA).

The dried pastes were then derivatised with a molar excess (80 µL) BSTFA $^{+}$ 1% TMCS (Sigma-Aldrich, USA) and 120 µL Ethyl Acetate (Sigma-Aldrich, USA) and heated to 70$^{o}$C shaking at 1000 rpm in a thermocycler for 1 hour. The paste was broken into fine particles using a sterile blunt ended instrument to maximise availability for derivatisation. The volume was topped up to 200 µL to replace that which had evaporated and then samples were then centrifuged at max g for 10 minutes. Supernatents were then transferred to chromacol glass vials with inserts and crimped shut (Thermo Fisher Scientific, USA) and analysed on GC-MS. The resultant concentrations were then multiplied by the dilution factor of 2.

### 2.7.2 Standard preparations

Initially standards of 2TMS derivatised 6-ACA were prepared by firstly derivatising 1ml at a concentration of 1g/L or 10g/L using 300 µL BSTFA $^{+}$ 1% TMCS in with 700 µL. Derivitisation was complete when the white solid of 6-ACA dissolved. This was then subsequently diluted in ethyl acetate to make a range of appropriate final concentrations. As indicated in the results section 5.2.2, there were two other peaks present which were the main molecules present in the standards. In order to circumvent this all standards were prepared as follows. A number of induced pQR1553 cells were inoculated into 100 mls LB broth in non-baffled flasks and induced as previously described. They were not fed any cyclohexanol. They were grown for 24 hours to ensure completed fermentation growth. 6-ACA standard was weighed and added to each culture vessel i.e. one flask corresponding to one concentration of 6-ACA, a second to smaller concentration and so on. It was dissolved and shaken for 5 minutes and 1 ml samples were taken, sonicated and dried in the same

manner as if the culture produced it themselves. This method produced a linear and repeatable curve and the trace lacking the presence of either of the previously two identified molecules. This method was also used for cyclohexylamine. The first method of bulk derivatising the standard was used for 6-hydroxyhexanoic acid and adipic acid due to the presence of only a single peak on the GC trace. All other standards were weighed and prepared to 1 g/L in 1 ml of ethyl acetate. Several standard concentrations were then prepared from the 1 g/L standard in ethyl acetate and transferred to chromacol glass vials with inserts and crimped shut (Thermo Fisher Scientific, USA) They were then analysed on GC-MS and a curve was made using the Chromeleon™ software. Cyclohexanol, cyclohexanone, cyclohexylamine and adipic acid were all sourced from Sigma-Aldrich, USA. 6-hydroxyhexanoic acid and 6-ACA were sourced from Alfa Aesar, USA.

## 2.8 Gas chromatography mass spectrometry

A Thermo Fisher Scientific Trace 1300 GC coupled with an ISQ QD single quadrupole mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA) was used to identify pathway intermediates. The column inlet was heated to 270$^{o}$C, MS transfer line to 300$^{o}$C and the ion source to 300$^{o}$C also. These temperatures were chosen so as to completely vaporise the sample as well as maintain the phase so as to not condence in the machine. The column was pre-washed with 2 cycles of Ethyl acetate followed by 1 µL of sample was injected into the column using a split ratio of 1:5. This ratio was predetermined after trial and error to achieve the best peaks and detectability available. Total run time was 11.2 minutes consisting of 1 minute at 85$^{o}$C followed by a 5$^{o}$C/min ramp to 100$^{o}$C then a 25$^{o}$C/min ramp to 280$^{o}$C resulting in a 11.2 minute run. This method was optimised using standards and cell culture samples. The MS detector was set to scan within a range of 10-300 amu. A Restek RXI-5 sil MS, 30 meter, 0.25 mm I.D and 0.5 mm O.D column was used and at times

both a 0.5 µm and a 0.25 µm film thickness was used (Restek, Sauderton, U.K.). An alternative column used was the Agilent HP-5ms Ultra inert, 0.25 mm I.D, 0.5 µm film thickness and a 30 m total length (Agilent technologies, USA). The column was conditioned when the baseline at the latter end of the run rose. This was done according to manufacturers' instructions. Retention times of the intermediates on a 0.5 µm thick film were as follows; Cyclohexanol (2.53 min), cylohexanone (2.63 min), ε-caprolactone (5.5 min), cyclohexylamine (6.5 min), 6-hydroxyhexanoic acid (7.73 min), 6-ACA (8.1 min) and Adipic acid (8.55 min). These times were slightly altered when a 0.25 µm column was used or when a different flow rate was chosen. In order to reduce burden and fouling of the ion source, detection times were often shortened to 9 minutes when all the pathway intermediates had been detected. This was to allow the higher molecular weight molecules to burn off and vent.

# Chapter 3: Assembling the *Acinetobacter calcoaceticus* NCIMB 9871 genes

# 3 Results

## 3.1 Isolating the gene fragments and creating homologous flanking sequences

Amplifying genes from genomic DNA is intrinsically more difficult than isolating from DNA libraries primarily because of the difference in size. Standard PCR primers are typically 18-25nt in length and should theoretically only bind to the specific sequence. However if the primer has significant homology with other unknown stretches of DNA, some primers may anneal to these instead of the target sequence and therefore amplifying off-target products. Owing to the size difference, it is likely that this will occur more frequently in genomic templates. Therefore in cases where the gene of interest has flanking regions which have a more ideal GC content (and therefore are more likely to be more specific primers) than the gene itself, it may be more preferable to use a nested approach to increase the yield of target DNA. This is the case with *Acinetobacter calcoaceticus* NCIMB 9871 genes of interest. There are short sections at both the 5' and 3' ends of the target genes in this pathway which have low GC contents, therefore a region outside of these genes was chosen leading to a target gene flanked by a a small portion of genomic, non coding DNA. In the first instance, two genomic DNA samples and one lambda phage clone (λ8.1; confirmed by Southern blotting previously by Prof. J. Ward to contain *ChnB* and so is likely to contain the pathway) were used in a PCR gradient amplification of *ChnA*. A DNA agarose gel confirmed the generation of a predicted 944bp fragment in both genomic samples but was absent in λ8.1 (Figure 3.1 1% agarose gel confirmation of an expected band at 944bp containing *ChnA*.).

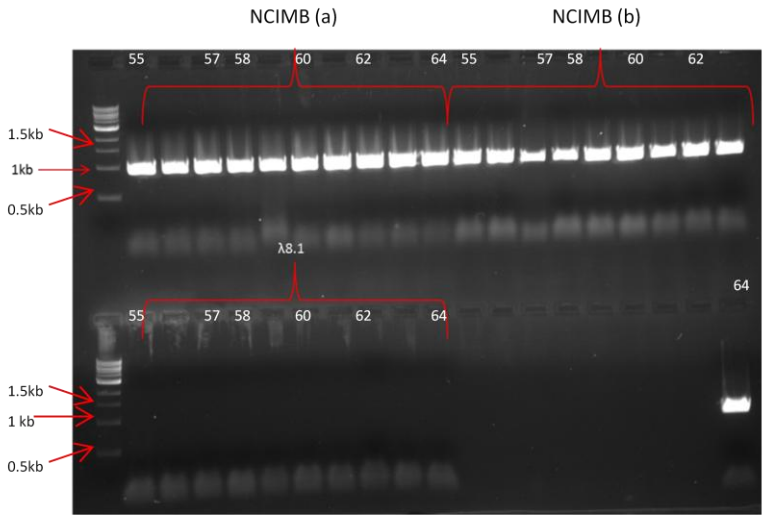



**Figure 3.1 1% agarose gel confirmation of an expected band at 944bp containing *ChnA*. The gene can be identified in both genomic samples however it appears to be absent in the lambda clone. Above the wells is the annealing temperature for the gradiant PCR.**

As can be seen in Figure 3.1, although *ChnA* amplified in both genomic samples, it failed in λ8.1. The clone had been positively identified as encoding *ChnB* by Southern blotting previously and so to confirm this a PCR reaction was performed on λ8.1 and NCIMB9871(a). This was confirmed as a fragment of length 1752bp on a 1% agarose gel (Figure 3.2).

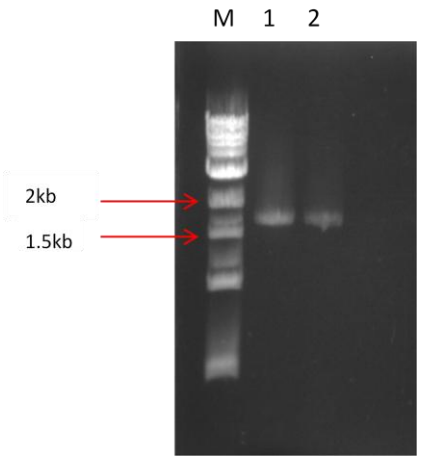



**Figure 3.2 1% agarose gel of PCR fragment containing *ChnB* from λ8.1 in lane 1 and NCIMB(a) in lane 2. The Marker lane is a 1kb NEB marker. Results show that the gene is present in both samples. Therefore the lambda clone is likely to be missing *ChnA* but does have *ChnB* which was the marker to confirm previously clone of the cluster was successful. Each gene is several thousand basepairs away from each other in the cluster.**

Given the incomplete pathway shown to be contained within λ8.1 all further PCR amplifications of the nested gene fragments were performed using NCIMB(a) genomic DNA (Figure 3.3).

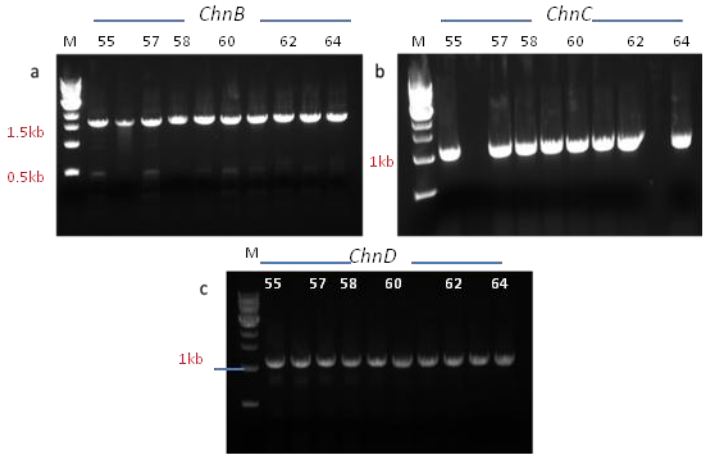



**Figure 3.3 Nested fragment gradient PCR amplifications. Off target effects can be seen at lower temperatures. a) *ChnB* (1752bp), b) *ChnC* (1040bp), c) *ChnD* (1137bp). Marker lane (M) is a 1kb ladder (NEB, USA). The temperature gradiant is displayed above each lane. All samples have amplified out successfully and with minimal off-target products**

Since there was significant yields at all temperatures in the gradient, the PCR samples from the individual reactions were pooled together and 30 µL of each fragment was ran on a 0.8% agarose gel and purified. The nested fragments were then used as templates to amplify the overlapping fragments with either their natural RBS (as identified by eye) or the RBS AGGAG at position -7 were amplified into the gene fragments to be used for assembly. The pET29a vector was also PCR amplified to contain overlaps for the 5' end of *ChnA* and 3' end of *ChnD*. Again, the fragment bands were excised from an agarose gel and purified ready for operon assembly. Significantly lower yields were obtained when purifying out the

pET29a vector owing to the difficulty in purifying larger fragments using column based separation.

## 3.2 Operon assembly

### 3.2.1 Gibson assembly as an operon construction tool

Owing to varying reports of the success of Gibson assembly, the method was first tested to see how efficiently the reaction assembled the 4 genes in the proposed pathway. Equimolar amounts of each fragment were incubated following the manufacturer's instructions at 50°C for 1 hour. The reaction mixture was visualised on a 1% Agarose gel (w/v).

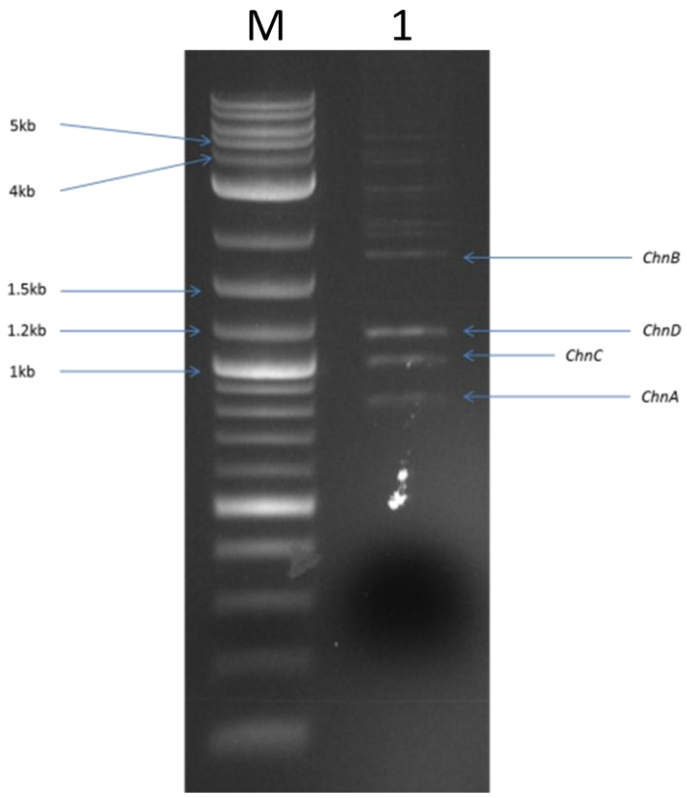



**Figure 3.4 Gibson assembly post 1 hour incubation on 1% agarose gel (lane 1). The remaining 4 gene parts are highlighted and several smaller bands of larger size can be identified indicating a small degree of assembly. The marker used was a 2-log ladder (NEB, USA)**

As can be seen in Figure 3.4, following the 1 hour incubation a significant amount of the genes have failed to incorporate into a single 4.4kb expected band and several faint bands of larger size can be seen. Note this is an assembly reaction and not an amplification so the DNA loaded is a very small amount. However, this indicates that the efficiency of Gibson assembly at assembling multiple fragments may be lacking. Given that two fragment assemblies have been reportedly more efficient, *ChnA* and *ChnB* were assembled in one reaction and *ChnC* and *ChnD* were assembled in a second, with the idea of then assembling the two larger fragments together in a two-step reaction. However, whilst *ChnC* and *ChnD* assembled with approximately 100% efficiency to create a ~ 2kb fragment, *ChnA* and *ChnB* completely failed to join (Figure 3.5). This image is poor quality however the reason for this that the image is overexposed so as to make visible the picomoles of DNA present.

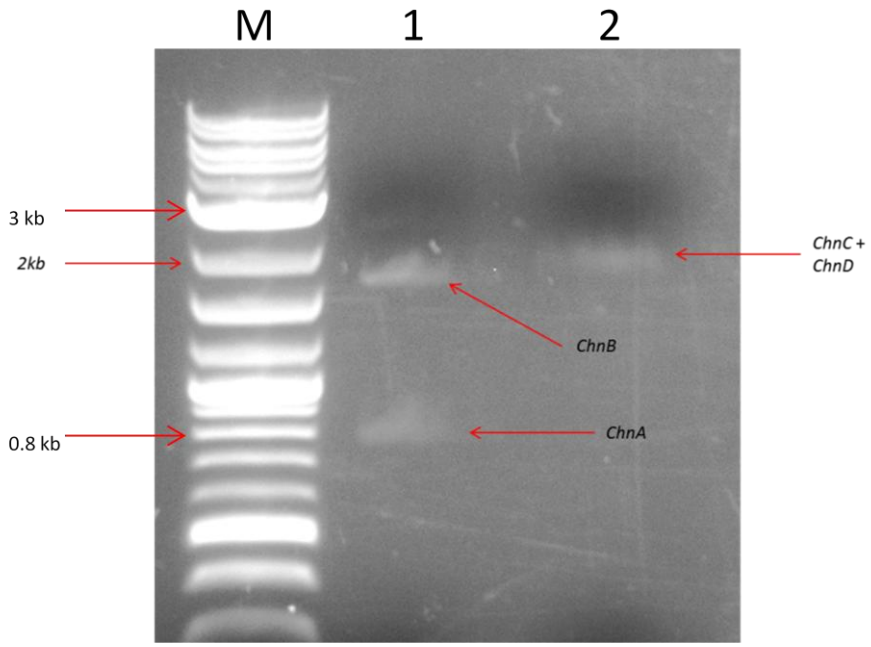



**Figure 3.5 Paired fragment assembly. A total of 0.5pmol of fragments were assembled in each reaction as stated by the manufacturers for a two part assembly. Lane 1 shows the lack of assembly between *ChnA* and *ChnB* whilst lane 2 shows the completed assembly between *ChnC* and *ChnD.* Marker (lane M) is a 2 log ladder, NEB**

In order to possibly increase the success of the 4 fragment assembly, an increase in the amount of time for the 50°C incubation step was compared to an increase in the amount incubation time plus increase in the amount of enzyme mix. In both cases the time was doubled to 2 hours and in the increased enzyme case, the volume of water used in previous reactions to bring the reaction volume to 20 µL was replaced with Gibson assembly master mix.

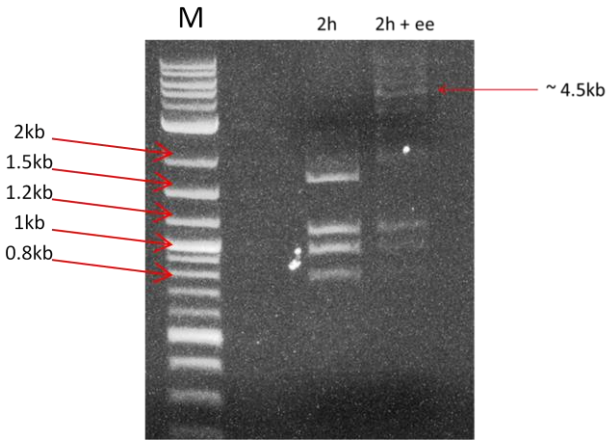



**Figure 3.6 Excess time is compared to excess enzyme in order to increase the proportion of each fragment that are assembled into an expexted 4.5kb stretch of DNA. It can be seen that with extended tiem and excess enzyme, there is less remaining DNA fragments compared to only excess time. There is also a band at 4.5kb which could be the expected assembly. Therefore, after following the protocol from the manufacturer, in order to optimise assembly, excess enzyme is required.**

Results of this experiment suggested that most likely an increase in the amount of enzyme, above what was recommended is needed as it greatly increased the efficiency of fragment assembly (See www.NEB.com for the gibson assembly protocol). As can be seen in Figure 3.6, as well as two or more fragment bands being present at around 4.5kb can be observed. However, less of the faint bands were seen in the 2 hour incubation compared to the 1 hour incubation suggesting significant experiment variation in efficiency.

With the observed inefficiencies in assembling 4 small fragments, the prospect of adding a 5$^{th}$, much larger fragment (*NdeI* and *NotI* double digested pET29(a) plasmid) of 5.23kb, to the reaction looked likely to be unsuccessful. With this in mind, a double digestion of the 4 fragment assembly was performed and ligated with the vector and subsequently transformed. The number of colonies that arose was excessively large indicating perhaps incomplete digestion of plasmid and transformation of circularised pET29(a) vector. One colonies plasmid was isolated and then digested singularly and double digests according to manufacturer's instructions to identify the 4.4kb operon from a total construct of 9.63kb. Figure 3.7 shows an undigested, *NotI, NdeI* and double digests and the agarose gel confirmed the absence of the two bands (4.4kb operon and 5.23kb vector). What can also be seen is there is a prominent band at just below 4kb, a faint band at just above 5kb and one above 10kb. This suggests a partial DNA digestion was loaded and that three forms of DNA can be seen. The total length of pET29(a) plasmid is 5.371kb and therefore the band lower than 4kb is possibly supercoiled DNA, the faintest band above 5kb is linearised (fully digested DNA) whilst the largest faint band is possibly nicked/relaxed circular. Therefore not only is there no operon but also the digestions haven't worked and are likely to require a longer digestion time e.g. 1$^{+}$ hours. Since neither *NdeI* or *NotI* have star activity, an overnight digestion is possible.

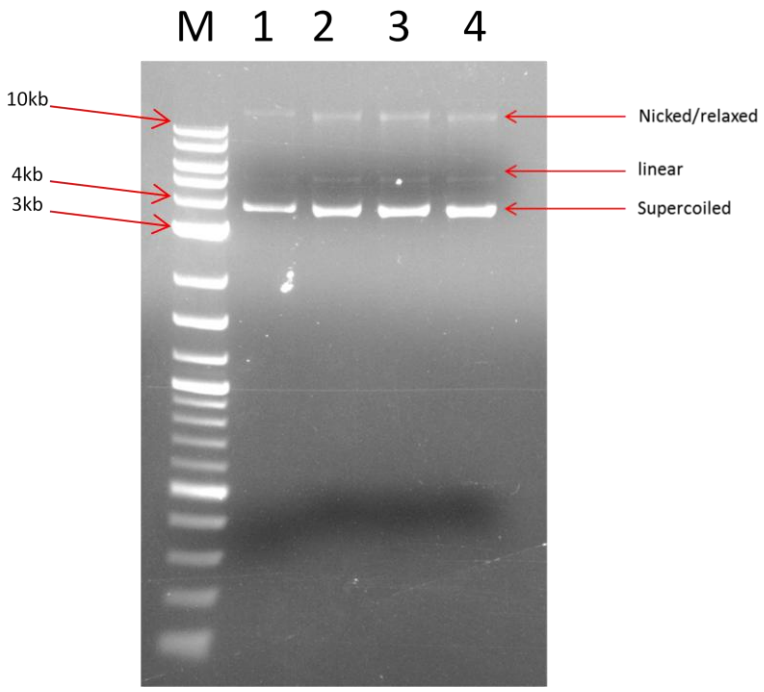



**Figure 3.7 A gel of undigestd (lane 1), *NotI* (lane 2), *NdeI* (lane 3) and *NotI* and *NdeI* digested (lane 4) of a plasmid miniprep from a colony from the Gibson assembly transformed cells. Three forms of DNA can be seen and no operon insert. Results suggest this length of time of digestion is insufficient as directed by the manufacturer (NEB, USA).**

Results indicate that Gibson assembly is an inefficient method of assembling this multi-fragment gene segment however to comment further would require a 5 fragment Gibson reaction and transformation and subsequent complete digestion of the minipreps. Future digestions with *NdeI* or *NotI* will also require a longer incubation period.

### 3.2.2 CPEC as an operon construction tool

Circular polymerase extension cloning (CPEC) is a novel method of DNA assembly requiring no exonucleases or ligases and therefore reactions can be performed with higher DNA concentrations in a thermocycler compared to proprietary kits. Since the fragments which were generated for Gibson assembly contained the desired overlaps, the same fragments were used in a CPEC reaction to assess if they would anneal as described by Quan and Tian (Quan and Tian, 2009). Initially only the 4 gene fragments were used in equimolar amounts

and visualised on an agarose gel. A single prominent band was observed with a small smear which could possibly indicate linear fragments of different sizes. However a much larger conversion into the 4.4kb fragment can be observed compared to Gibson assembly of the same fragments (Figure 3.8).

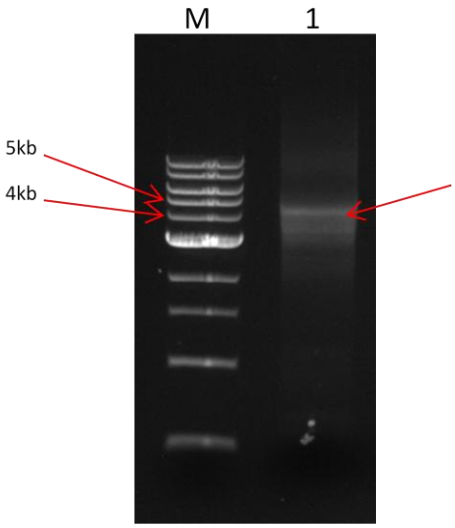



**Figure 3.8 CPEC reaction of the 4 gene fragments used to create operon A showing a single prominent band at ~4.4kb from . Minor smearing can be observed however these are thought to be linear fragments. A much larger yield of joined fragment is seen compared to Gibson assembly. The marker lane (M) is a 1kb ladder (NEB,USA), 5µL of the CPEC reaction was loaded onto a gel (lane 1) and a single band can be seen between 4 and 5 kb. Smearing either side could indicate varying length DNA.**

This showed the potential for CPEC to be a plausible method of operon construction. The pET29a vector was amplified with a high fidelity polymerase to contain a 5' homologous sequence for *ChnA* and a 3' homologous sequence to *ChnD*. The sizes and properties of the homologous regions between the fragments are summarised in Table 3.1.

| Section | Length of homology (bp) | Tm (°C) |
|---|---|---|
| **Vector-*ChnA*** | 49 | 60 |
| ***ChnA*-*ChnB*** | 43 | 65 |
| ***ChnB*-*ChnC*** | 44 | 65.5 |
| ***ChnC*-*ChnD*** | 55 | 69.8 |
| ***ChnD*-Vector** | 49 | 73.8 |

**Table 3.1. Summary of the homologous overlaps between the fragments and the Tm of the stretch of DNA**

A second CPEC cycle was then performed using all five fragments (including amplified pET29a vector incorporating homologous sequences) and in both operon A and operon B a small sample was visualised on a 0.8% agarose gel as well as directly transformed. An agarose gel of the 5 fragment CPEC reaction showed an unusual circular DNA band which failed to move past the well (Figure 3.9). This could be possible concatenation and can be seen in the high cycle number CPEC reactions (30 cycles) shown in figure 2 of Bryksin and Matsumura (Bryksin and Matsumura, 2010). Despite this, the transformation yielded many colonies. Several colonies were then chosen for overnight growth and miniprepped for restriction digestion in order to ascertain which colonies contained the operon in the correct alignment.

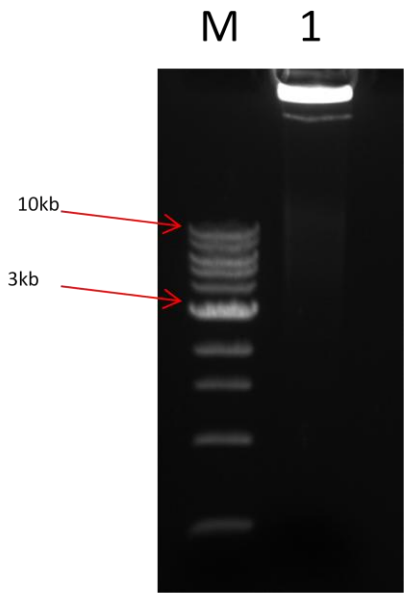



**Figure 3.9 0.8% agarose gel of the 5 fragment CPEC reaction for operon B shows a thick band which moved little from the well. Several reasons for this include an entanglement of DNA when assembling, contatenation of the DNA orprotein impurities binding the DNA. Interestingly this effect is only seen when the product formed is circular. It is not seen when the product is linear (lane 1). Marker (lane M) is a 1kb ladder (NEB, USA).**

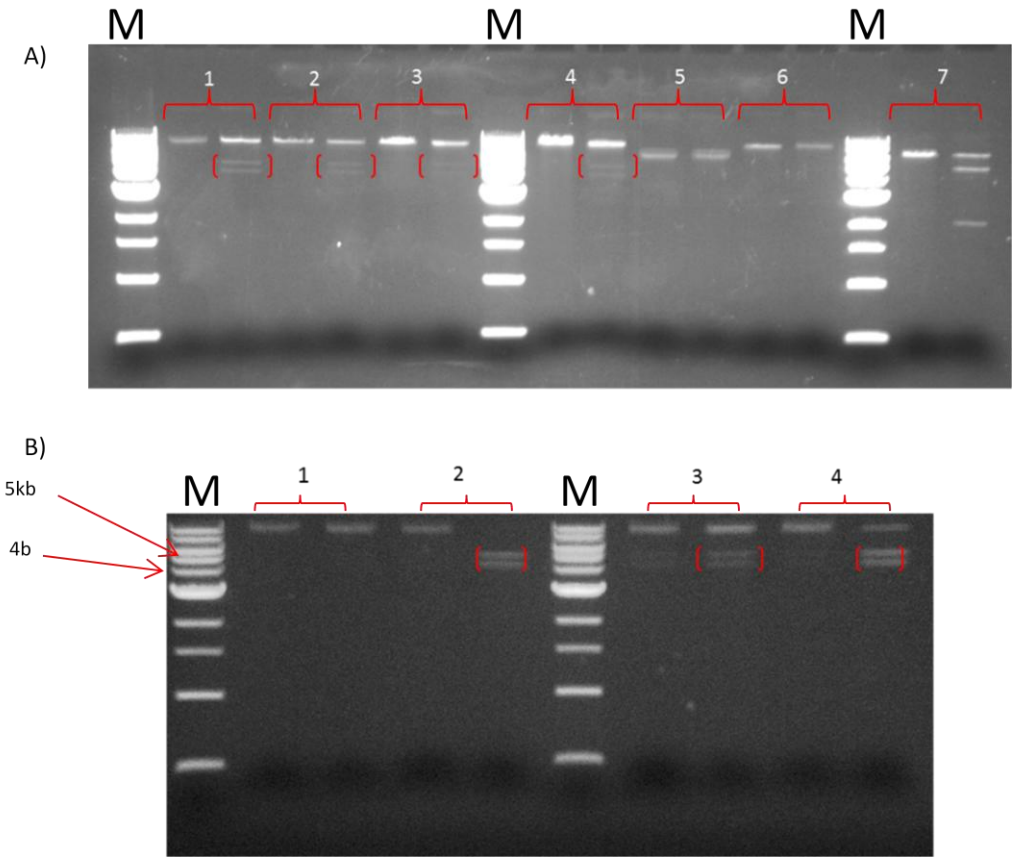



**Figure 3.10 Operon A (native RBS) (Top - A)): an agarose gel of 7 miniprepped colonies digested singularly with *NdeI* (first lane of each sample) and a second with *NdeI* and *NotI* (second lane of each sample). Bottom (B)): miniprepped plasmids 1- 4 from above comparing 1 hour *NdeI* and *NotI* (left lane of each sample) to 16 hours (right lane of each sample). Expected bands with the double digestion is 4.4kb operon and 5.2kb backbone. It can be seen a 16 hour digestion greatly improved digestions. Colony 1 failed to digest but contains two fragments ( right lane of A) sample 1). 4/7 colonies therefore exhibited the correct digestion. Lane M denotes the marker lane using a 1 kb ladder (NEB,USA).**

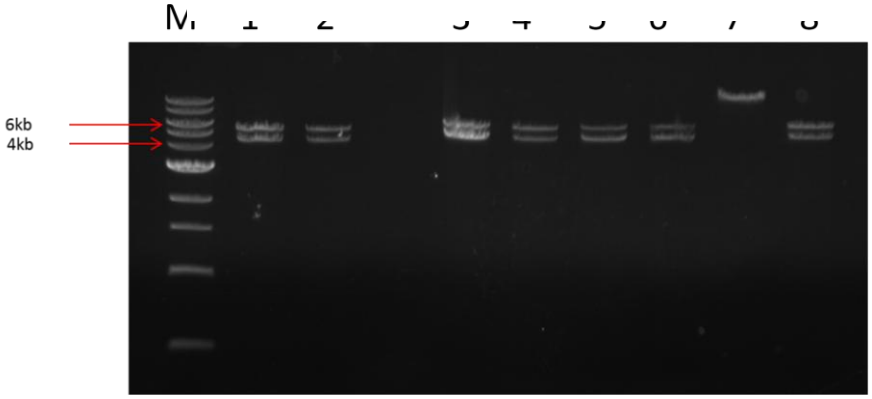



**Figure 3.11 8 colonies from the operon B assembly (lanes 1-8) were miniprepped and double digested with *NdeI* and *NotI* for 16 hours. Highlighted are the 4kb and 6kb DNA ladder bands in the marker (M) lane. The two digested bands fit between these as expected. Only colony 7 failed to digest as expected.**

Out of the 7 colony DNA samples of operon A digested, 4 digested in the correct manner whereas operon B, 7 out of the 8 digested as expected with *NotI* and *NdeI*. The plasmid size is 9.632kb and when double digested produce two bands of 4.4kb and 5.2kb, i.e. the operon and backbone respectively. Given that each gene is of similar size, digestion of the colonies with only the enzymes at the beginning and at the end does not mean that there is only one copy of each or that they're in the correct orientation. Theoretically the fragments should anneal to the specific partners however even with proprietary kits like Gibson assembly, misallgnments and incorrect assemblies occur (figure 4 on: https://www.neb.com/products/e5520-nebuilder-hifi-dna-assembly-cloning-kit). Therefore each gene has to be digested with a specific enzyme which will identify any errors in assembly. *PstI* cuts once only through *ChnA* and so by digesting this with *NotI*, two bands can be identified of sizes 4151bp and 5481bp (Figure 3.12). If only one band appears then *ChnA* is absent. *HindIII* cuts nowhere on the pET29a plasmid and only through *ChnB* and *ChnD*, therefore if two bands are shown, one at 2451bp and another at 7181bp then both genes are present in the correct orientation (Figure 3.13). *AccI* cuts once through pET29a and once through *ChnC* therefore fragments of 5857bp and 3775bp would be expected (Figure 3.12).

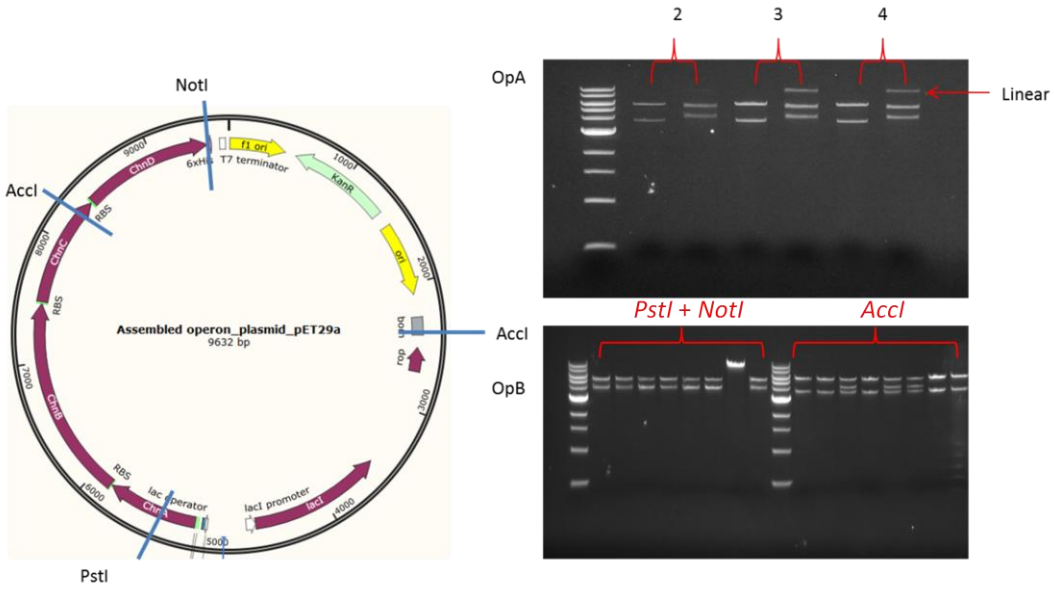



**Figure 3.12 Left; a schematic of the restriction map digestions. Right top gel; three colonies from operon A digested with *AccI* in the first lane of each sample and *PstI+ NotI* in the second lane of each sample. Right bottom gel; 8 colony minipreps from operon B assembly are digested with *PstI +NotI* in the first 8 and *AccI* in the second 8 samples. All three colonies in Operon A digested as expected whilst, although 7/8 contained *ChnA*, 3/8 showed irregularities in *ChnC* in the case of the operon B digests.**

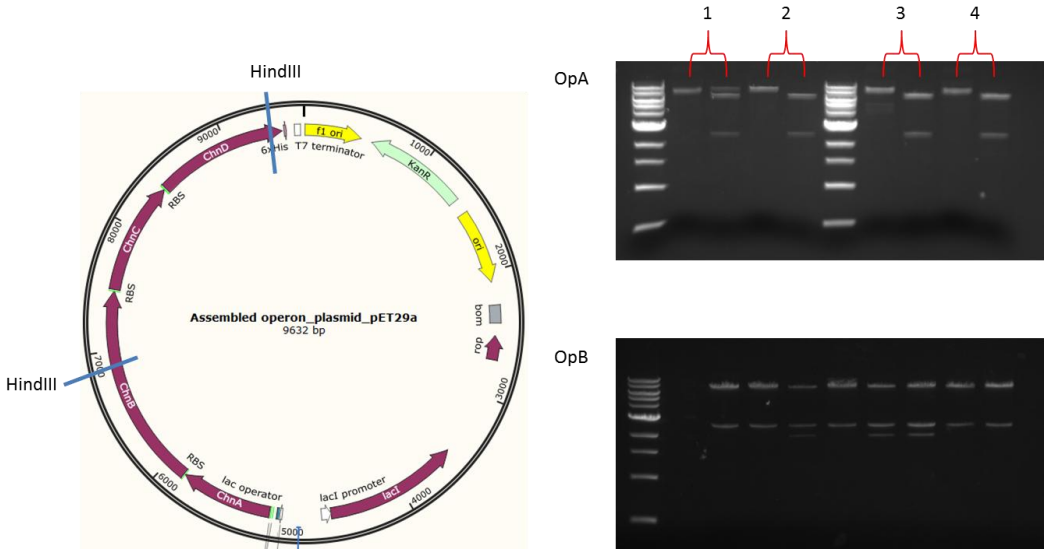



**Figure 3.13 Top; operon A, 4 colonies each with a linearised vector in the left lane (digest with *NotI*) and a *HindIII* digestion in the right of each. Only colony 1 of operon A looks to have some partitally digested plasmid with a single cut from *HindIII*. Bottom; *HindIII* digest of all 8 plasmid DNA's from the 8 colonies picked. Clone 3, 5 and 7 digested in the same incorrect manner.**

*N.b. Note that the only difference between the two clones is the RBS and so it is not possible to tell by looking at gel images which operon is which since there is only a few bases difference between the two.*

Taken together it is likely that 4/7 clones of operon A construction were assembled in the correct formation however only 3 of the DNA samples can be confirmed due to errors when preparing the experiments meant samples ran out. However in the case of operon B only 4/8 digested as expected.

### 3.2.3 Operon expression and sequencing

Once the plasmids with the correct assembly were identified, they were transformed into the pET29a expression strain of *E.coli* (BL21(DE3)). Overnight cultures were miniprepped and a sample stored as a glycerol stock. A plasmid sample of pQR1548 (operon B) and pQR1543 (operon A) was sent for sequencing of the beginning and ending operon regions using T7 primers. Sequencing confirmed *ChnA* and *ChnD* had correctly assembled as predicted. An example of the sequence readout is shown below (Figure 3.14). The incorporation of the *Nde*I site at the expected place as well as the RBS sequence was confirmed. The forward read out was 349bp and the reverse was 1028bp long and was also correct confirming zero mutations.

ACTTTAAGAAGGAGATATACATATGAGGAGAAAATTATGTCAAATAAATTCAA

**Figure 3.14 Segment of the sequence of Operon B revealed correct assembly as predicted. Blue: pET29a, Green: *Nde*I site, yellow: *ChnA* RBS and coding region. This is an example of the sequencing read out and note that the ends of the fragment on both operaon versions were confirmed as correct.**

Overnight cultures of each clone were then used to inoculate a shake flask and were induced with IPTG at $OD_{600}$ of 0.6-0.7 (see materials and methods; 2.3.1). A non-induced flask inoculated with the same overnight growth was used as control. Samples were taken periodically and a cell pellet collected for SDS-PAGE verification of soluble protein

expression. The SDS-PAGE (Figure 3.15) confirmed the presence of 4 separate and over-expressed protein bands at the expected molecular weights (*ChnA* = 25.7kDa, *ChnB* = 60.9kDa, *ChnC* = 33.2 kDa, *ChnD* = 37.2kDa). Although a negative control was not included for operon B, expression of 4 clear and distinct protein bands with the same patterning as the operon A can be seen. Interestingly *ChnB* expression cannot be seen at 23 hours. According to ExPASy ProtParam, *Chn*B has the lowest instability index and should therefore be the most stable however this has previously been noted by Chen *et al* (Chen, Peoples and Walsh, 1988). In accordance with their work, the protein bands looked most intense at 4 hours and were less so at 23 hours.

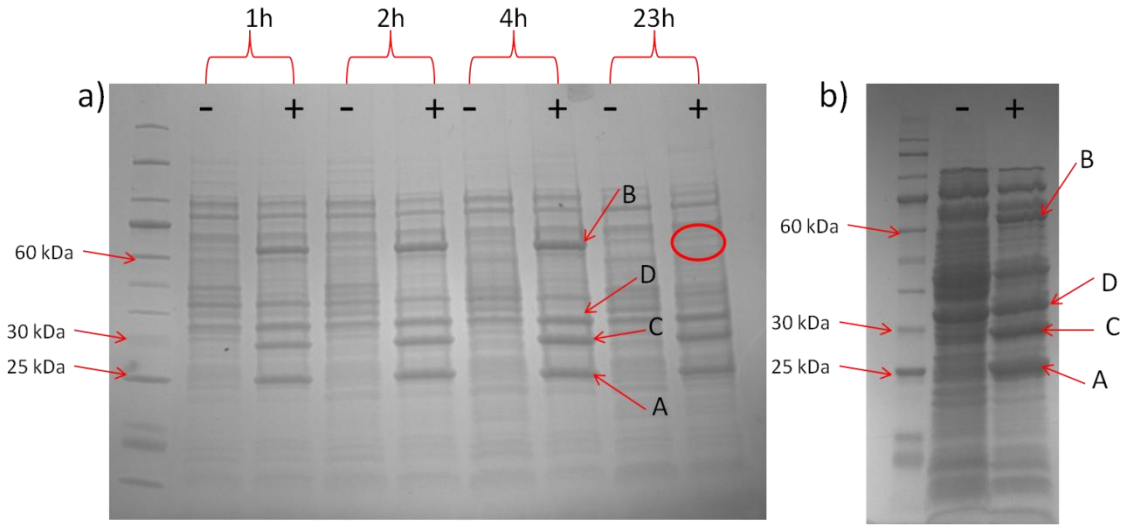



**Figure 3.15 SDS-PAGE of pQR1543 and pQR1548 (operon A (a) and B (b)) confirms the expression of the clone genes. *ChnB* is no longer present at 23 hours (circled) suggesting it is vulnerable to proteolytic cleavage and removal. Both operons are confirmed to have been constructed as expected by partial sequencing and express all 4 proteins intracellularly. By eye, it could be suggested that operon B expressed slightly more than operon A of *chnA, C* and *D* however possibly slightly less *chnB*.**

# Chapter 4: Individual enzyme cloning, expression and purification

# 4 Results

## 4.1 Cloning

In order to test if the enzymes the cells are expressing are in fact active, each enzyme has been cloned individually and modified by removal of the stop codon to incorporate a His-tag for nickel affinity purification. The reason for this is that detecting activity using singly expressed enzyme is easier than if all were combined and metabolic analysis was done e.g. by GC-MS. In the first instance *ChnA* was cloned into pET29a using standard restriction cloning techniques. Figure 4.1 shows the PCR amplification of *ChnA, ChnB, ChnC* and *ChnD* encoding *NotI* and *NdeI* flanking restriction sites.

The stop codon was removed to use the pET29a encoded His-tag. This was then digested along with the pET29a vector, ligated to each other and subsequently transformed into 10β cells. An overnight culture of a colony was then miniprepped and the plasmid subsequently double-digested with *Not*I and *Nde*I to confirm the incorporation of the *Acinetobacter* genes into the pET29a vector (Figure 4.2).

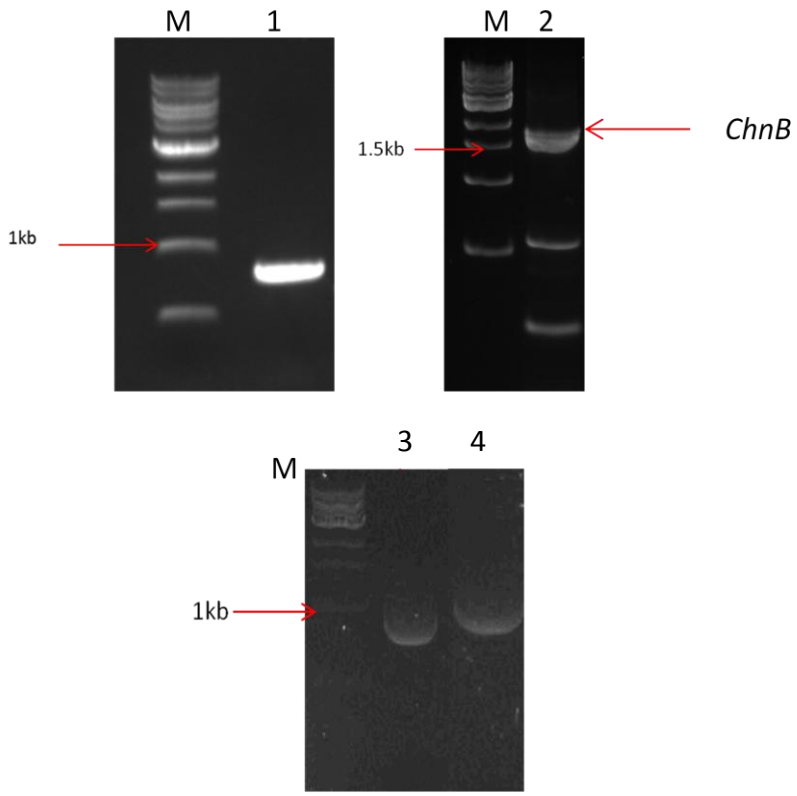



**Figure 4.1 Agarose gel to show PCR amplifications of *ChnA-D* incorporating individual restriction sites for restriction cloning. Lane 1 is *chnA* (744bp), Lane 2 shows c*hnB*(1632bp), lane 3 shows *ChnC* (903bp) and lane 4 shows *ChnD* (1059bp) were all successfully amplified. Only *ChnB* showed significant off target effects. The marker lane is a 1kb ladder (NEB, USA)**

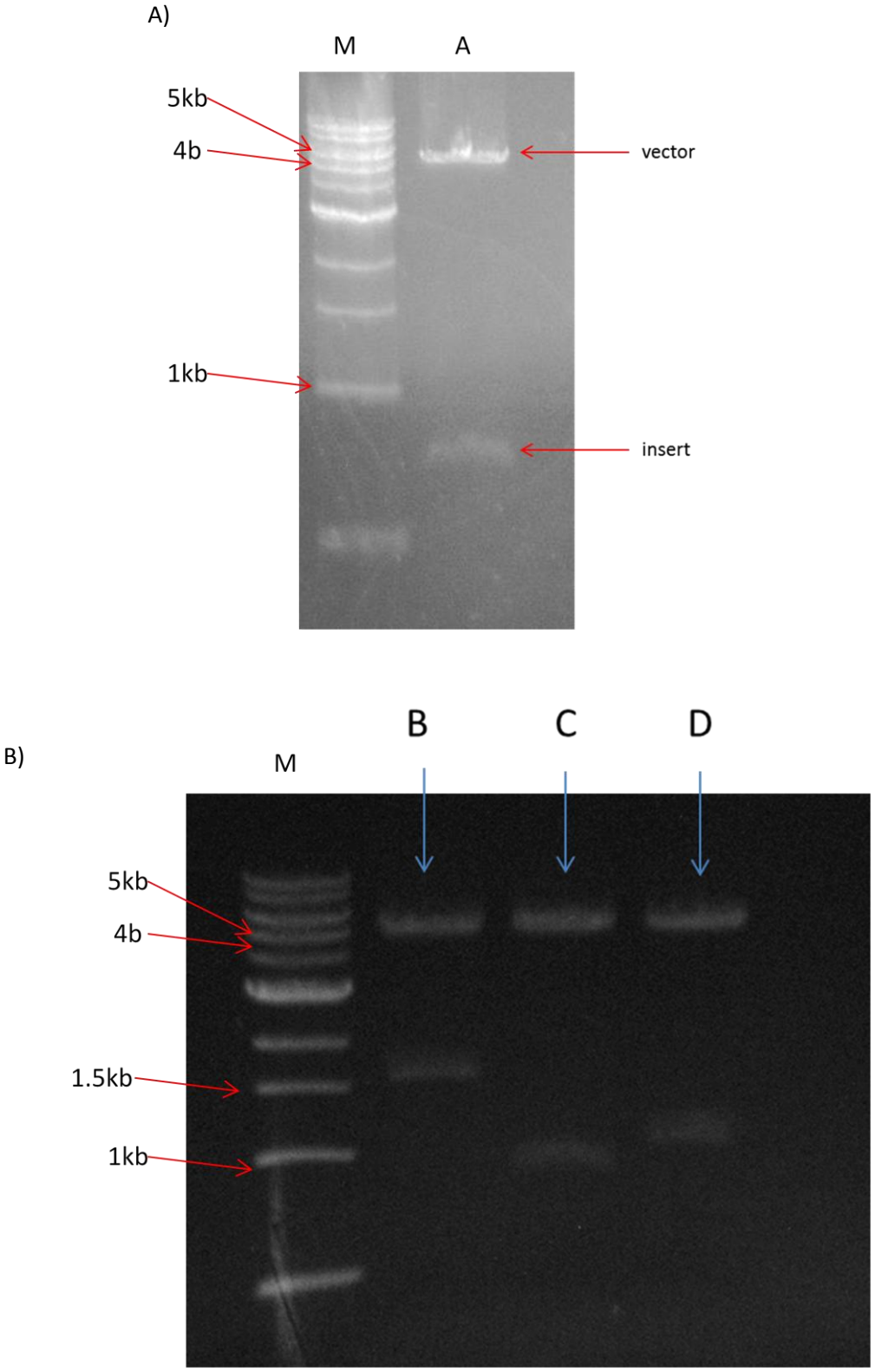



**Figure 4.2 Colony plasmid minipreps double digestion with *NdeI* and *NotI* confirms the insertion of *ChnA-D* into the pET29a vector. A) shows the correct digestion of *chnA* (Backbone in all is 5.2kb), *chnA* is 744bp) whilst B) shows the correct digestion of *chnB-D*(lane B, C and D; B is1632bp, C is 903 bp and D is 1059bp)). The ladder (lane M) is a 1 kb ladder (NEB,USA).**

## 4.2 Expression

The miniprep DNA was then transformed into the BL21 (DE3) *E.coli* expression strain and glycerol stocks of this were also made and renamed p*chnX* where *X* is the gene letter. A shake flask culture of the plasmid carrying clones was then performed and induced/non-induced (control) samples were taken. SDS-PAGE analysis confirmed an over-expressed band at the expected molecular weight of 25.7 kDa (p*chnA*) (Figure 4.3) 60.9 kDa (p*chnB*), 33.17 kDa (p*chnC*) and 37.19 (p*chnD*) (Figure 4.4) . This clone uses the RBS encoded on the pET29a plasmid which is presumed to be a more optimal RBS as well as the fact the cell is only over-expressing one protein compared to four and why there is significantly less of each protein expressed in either operon.

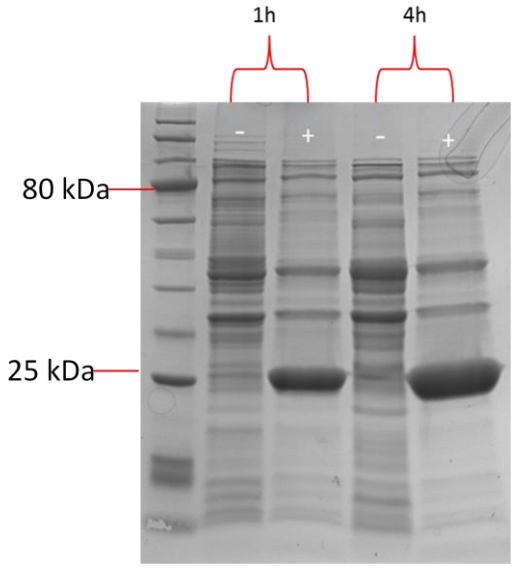



**Figure 4.3 SDS PAGE varification of p*ChnA* expression of cyclohexanol dehydrogenase in BL21 (DE3) cells. A prominent band can be seen at 25 kDa which intensifies with time**

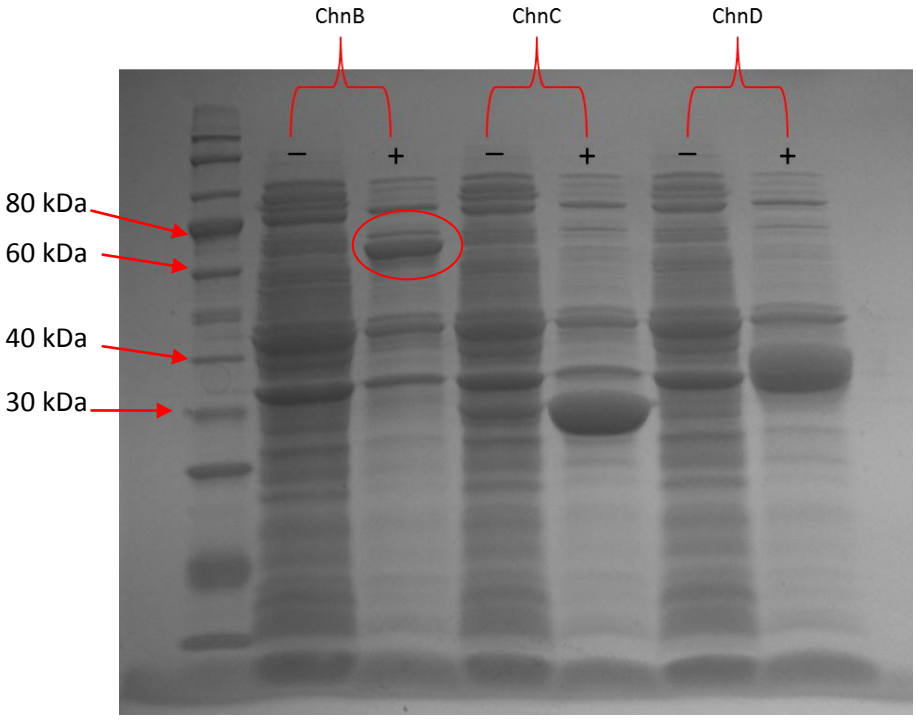



**Figure 4.4 An SDS-PAGE confirms the expression of the cloned p*chnA-D* plasmids in *E.coli* BL21 (DE3). Lane marked *ChnB* shows cyclohexanone monoxygenase circled. Expression of this CHMO appears to be substantially less than the other enzymes.**

## 4.3 Purification

Nickel affinity chromatography was then used to purify the individually expressed His-tagged proteins. Firstly not all of the protein was found to have been bound during column loading because as can be seen in Figure 4.5, Figure 4.6, Figure 4.7 and Figure 4.8 the flow through fractions contain a relatively large amount of protein. Secondly wash 1 removes a large proportion of host cell proteins but also some weakly bound his-tagged protein however by wash 6 most of it has been removed. After purification the different elution fractions were pooled and concentrated using a Vivaspin 20 concentrator followed by buffer exchange into 100mM phosphate buffer pH 7.5. To determine protein concentration, a 1 µL sample was analysed on a Nanopdrop at 280nm using an extinction coefficient which was calculated using the ExPASy ProtParam online programme. The concentrations are shown in Table 4.1 Although this is not a particularly high concentration, a relatively large amount of protein is lost in the buffer exchange and in the flow through. However, the concentration is sufficient for use in assays to identify activity. In addition it is unlikely that any residual co-eluted protein would have any influence in reactions and so it is considered to be sufficiently pure for activity assays.

| | |
|---|---|
| Cyclohexanol dehydrogenase | 0.385 mg/ml |
| Cyclohexanone monoxygenase | 77 µg/ml |
| Caprolactone hydrolase | 0.147 mg/ml |
| 6-hydroxyhexanoate dehydrogenase | 0.086 mg/ml |

**Table 4.1 Table of the final concentrations of each purified enzyme stored in 100 mM phosphate buffer at pH 7.5. Extinction coefficients were calculated from the ExPASy ProtParam programme. Aliquots were then flash frozen in liquid nitrogen and stored at -80°C.**

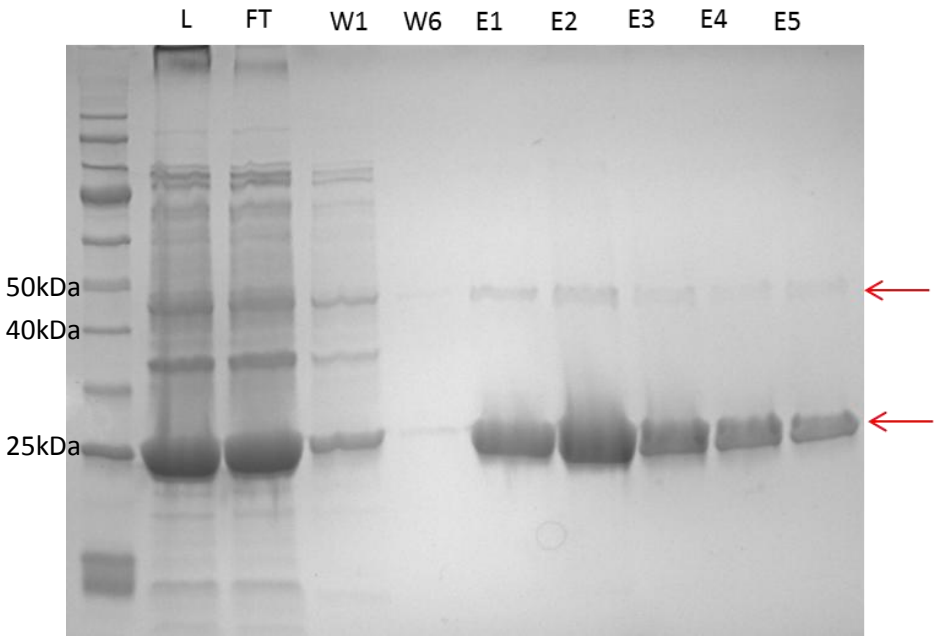



**Figure 4.5 SDS PAGE of the course of the protein purification. L (cell lysate), FT (flow through), W1/6(wash 1/6), E1-5(elution 1-5). Bottom arrow shows a prominent band at aproximately 25kDa corresponding to cylohexanol dehydrogenase and a second, fainter band at around 45kDa also co-eluted.**

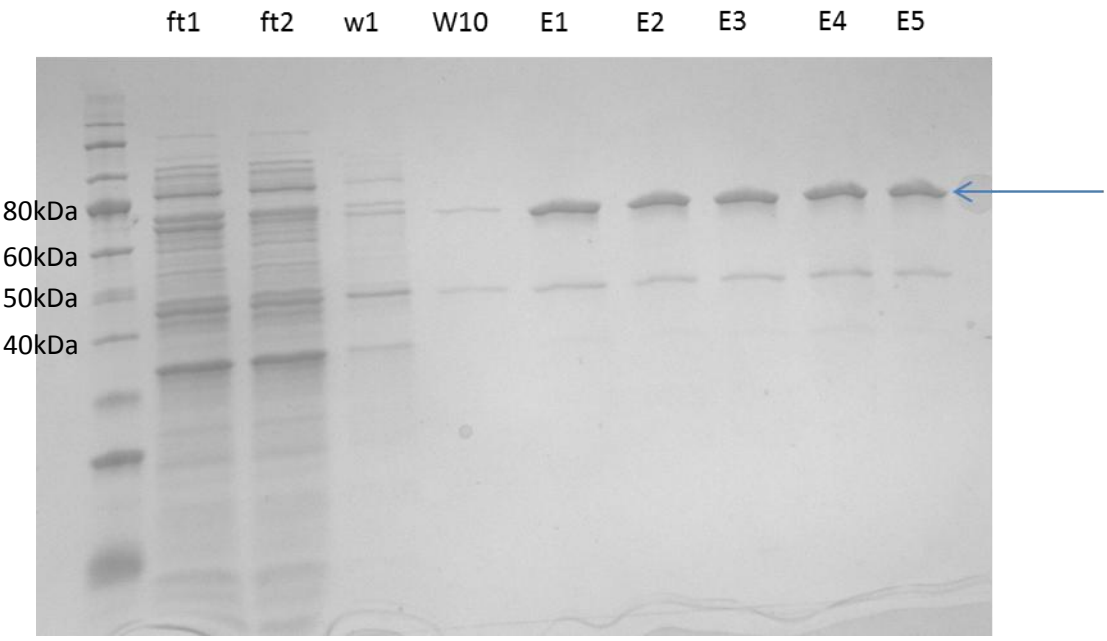



**Figure 4.6 an SDS-PAGE of the course of the purification of cyclohexanone dehydrogenase. Note the identical co-eluted band at 45 kDa similar to other protein purifications. Elution looks to be steady throughout the 5 elution steps. CL (cell lysate). FT (flow through), W1 (wash step 1), W10 (wash step 10), E1-3 (Eluted fractions 1-3).**

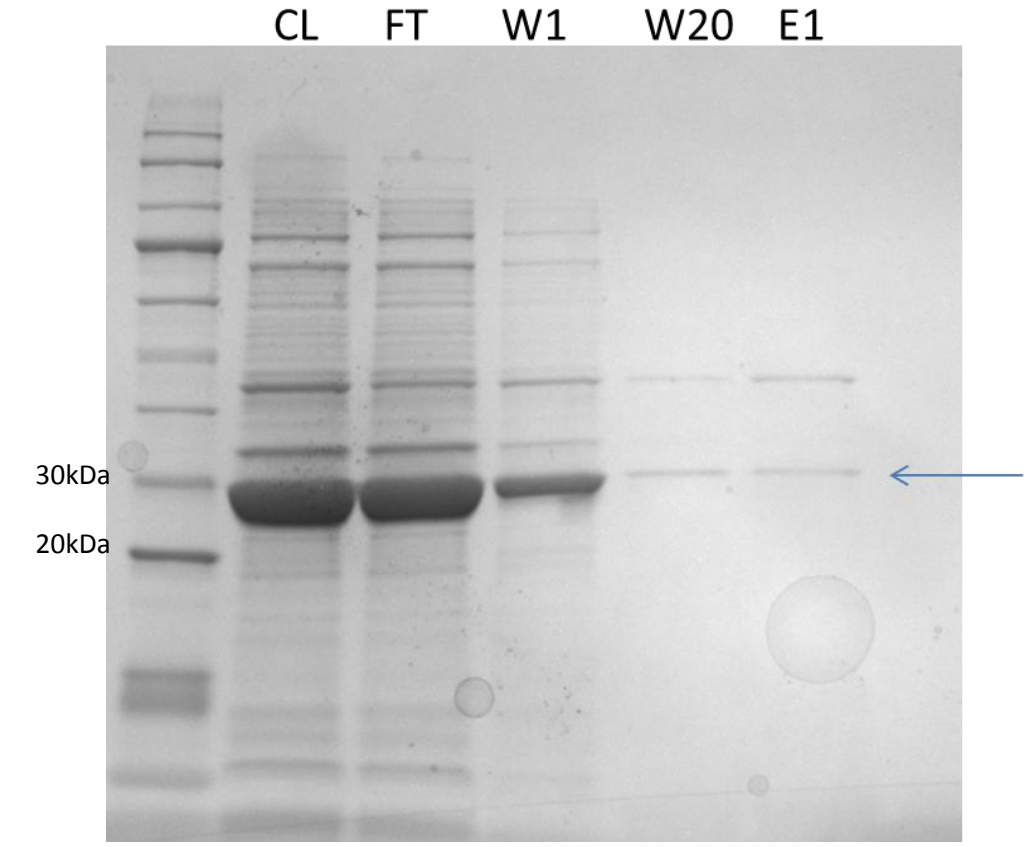



**Figure 4.7 An SDS-PAGE of the course of purification of the lactonase. Significant protein loss can be seen in the flow through and the first column volume wash. Very little lactonase remains in th eluted fraction (indicated by the blue arrow). CL (cell lysate). FT (flow through), W1 (wash step 1), W10 (wash step 10), E1-3 (Eluted fractions 1-3).**

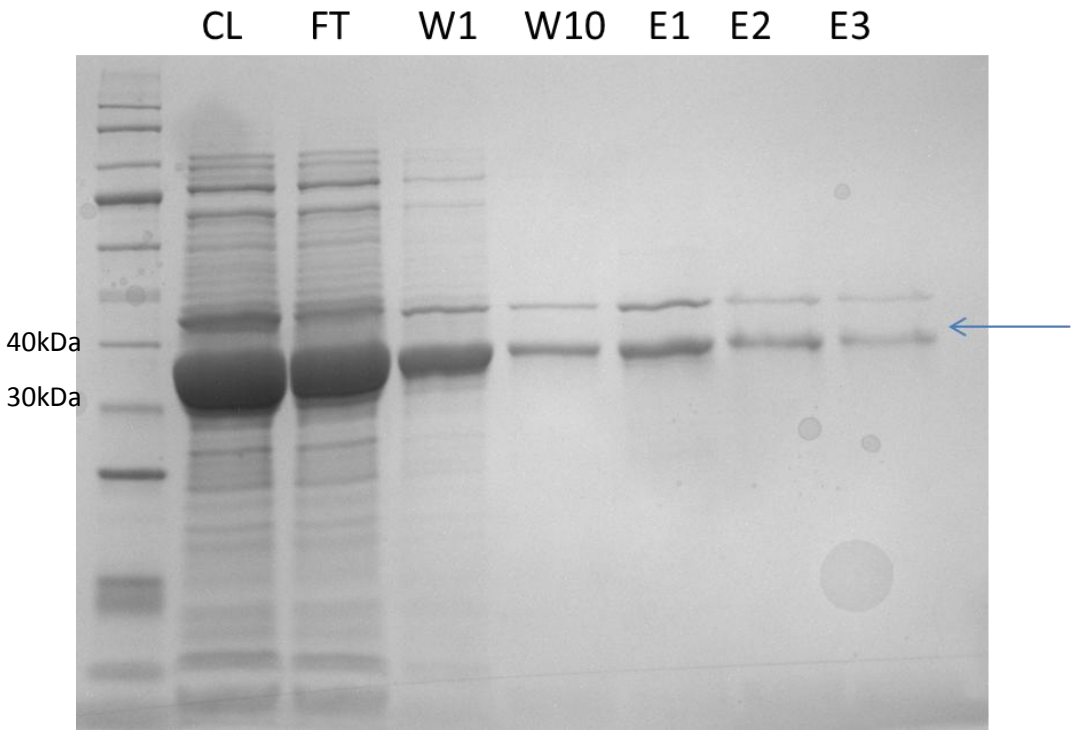



**Figure 4.8 An SDS PAGE showing the purification of 6-hydroxyhexanoate dehydrogenase. Similar to the lactonase in Figure 4.7, very little protein remains in the eluted fractions. CL (cell lysate). FT (flow through), W1 (wash step 1), W10 (wash step 10), E1-3 (Eluted fractions 1-3).**

## 4.4 Assaying for activity

Activity of enzymes was indicated by following absorbance changes of the required co-factors at 340nm. The co-factors being either the oxidation of NADH or the reduction of NAD(P). The aim of these absorbance assays is not to determine the absolute activity since the behaviour is likely to be markedly different *in vivo* and would provide little useful information, rather, if there is any activity at all it would likely indicate a functioning enzyme. Lack of functionality would have subsequently led to searching for alternative enzymes or directed evolutionary approaches with aims to increase to notIceable levels. A serial dilution of substrate concentrations were used with a final molarity of 50mM, 5mM and 0.5mM in all *Acinetobacter* enzyme assays. Since these experiments were purely speculative, a wide range of substrate concentrations were used. A final concentration of 0.5mM co-factor was used and the reactions were followed immediately for approximately 20 minutes or until the curves plateaued. 20µL of cell lysate or purified protein was used and an empty pET29a transformed *E.coli* BL21 (DE3) cell lysate or zero enzyme wells were used as a control.

### 4.4.1 Cyclohexanol dehydrogenase

Cyclohexanol dehydrogenase, encoded on gene *ChnA* (EC1.1.1.245), is an ezyme with low substrate specificity and also possibly thought to catalyse the reverse direction from a cyclic-ketone to cyclic-alcohols. Although very little work has been reported on this enzyme, indications suggest it is also capable of oxidising alicyclic diols and conversely reducing alicyclic diketones. Reaction rates have been reported to be between 0.8 µmol cyclohexanol · $min^{-1}$ · $mg^{-1}$ protein at pH7 up to 2.5 µmol · $min^{-1}$ · $mg^{-1}$ protein at pH10 with increased reverse reaction rates with a more acidic pH environment (Dangel *et al.*, 1989). Results using both purified protein and cell lysate suggest is possesses a high substrate

tolerace and a good activity. It appeared that it is able to maintain an excellent reaction rate at cyclohexanol concentrations as high as 50 mM. Higher concentrations were not tested however no substrate inhibition can be identified. Lower substrate concentrations reduced the reaction rate however this was not quantified (Figure 4.9; graph produced using cell lysates as an enzyme source)

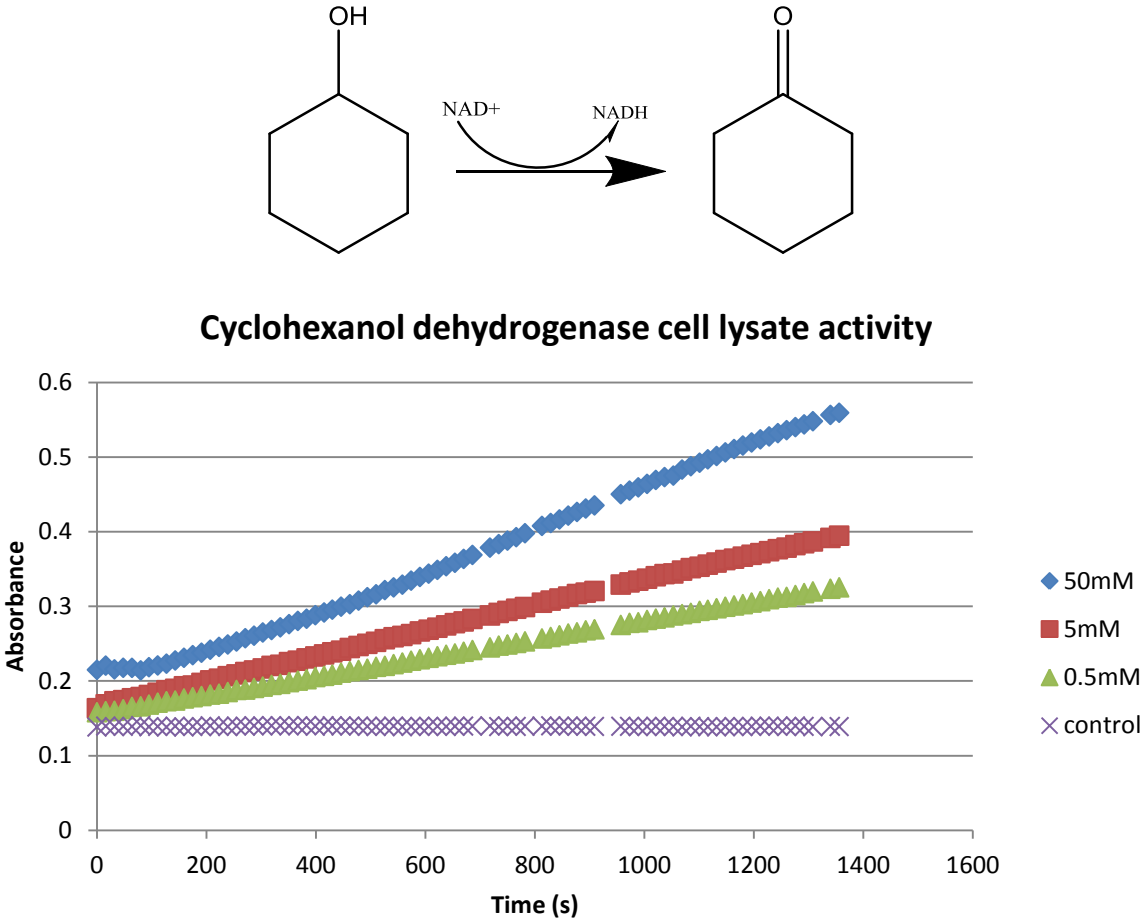
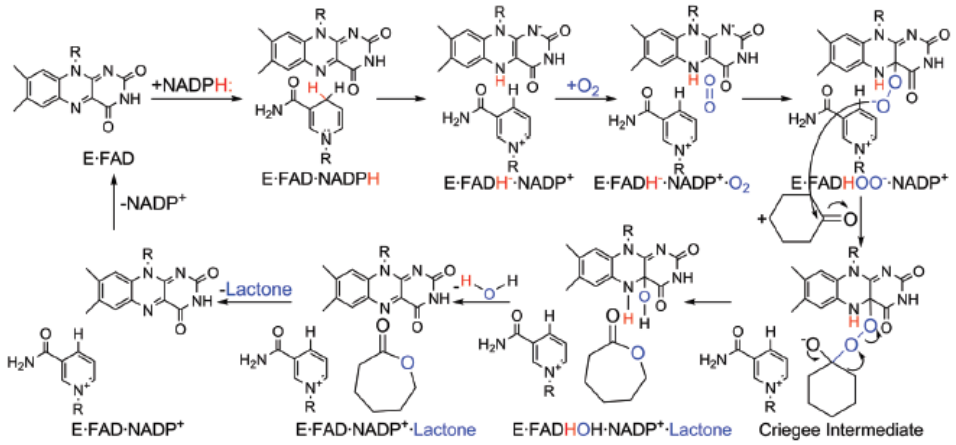



**Figure 4.9 The absorbance reaction at several concentrations of cyclohexanol using protein from a cell lysate (see section 2.4). 20µL of cell lysate was used in a 200 µL reaction in 100 mM phosphate buffer (pH 7.5). 0.5 mM $NAD^+$ was added, the reaction was followed at a wavelength of 340 nm. Note that this experiment was not aiming to quantify the activity but merely look to see if any is present and how it compares are different substrate concentrations. Therefore no kinetic parameters are reported. A schematic of the reaction is above the graph.**

### 4.4.2 Cyclohexanone monoxygenase

Cylohexanone monoxygenase (hereafter referred to as CHMO, EC 1.14.13.22) is a member of the Baeyer-Villieger monooxygenase family belonging to the larger super family of FAD bound flavoenzymes. This group of enzymes have been extensively studied with CHMO being amongst the most well studied. Much hype around this enzyme was generated due to it characteristically being able to produce lactones and esters from cyclic ketones and aromatic aldehydes producing water as the only by-product. A broad substrate specificity and excellent enantiomeric excess values are ideal for use in large scale chemical manufacture such as the production of new pharmaceuticals (Sheng, Ballou and Massey, 2001). The mechanism of action of CHMO is as follows; the non-covalently bound FAD molecule is reduced by NADPH and then reacts with molecular oxygen to form a C4A-paraoxyflavin intermediate. This intermediate then acts as a nucleophile and attacks the carbonyl carbon of the ketone forming a Criegee intermediate which then cleaves giving a lactone and a hydroxyflavin product. Water is then released and the hydroxyflavin is oxidised back to FAD. $NADP^+$ then leaves the complex (Mirza *et al.*, 2009).




**4.10 The mechansim for the CHMO to convert cyclohexanone into a lactone using enzyme bound FAD. The figure is copied directly from Mirza *et al.*, 2009**

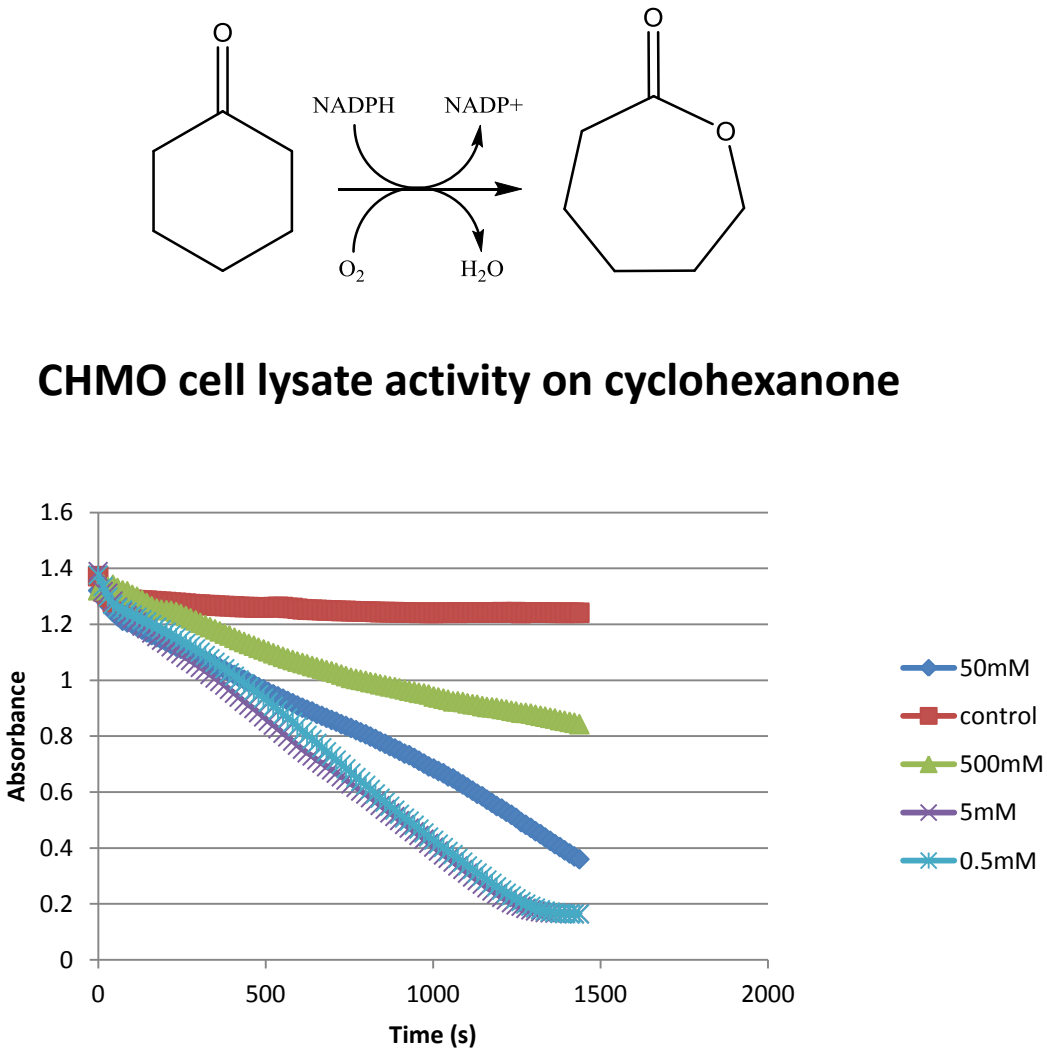



**Figure 4.11 Top is a schematic of the reaction. Bottom, the absorbance reaction of the CHMO using different concentrations of cyclohexanone. Substrate inhibition is shown in agreement with the literature. See material and methods for details on the absorbance reaction (2.4). In brief, 20 µL of cell lysate was added to a total reaction volume of 200 µL in 100 mM phosphate buffer (pH 7.5) with 0.5 mM NADPH and the absorbance was measured and followed at 340nm. As with all individual enzyme reactions, no reaction rate was calculated. Instead any indications of the enzyme behaviour including response to different substrate concentrations and any inhibitions was investigated. A schematic of the reaction is above the graph.**

After cloning into pET29a and fusing with an N-terminal His-Tag, expression levels were relatively low. Purification yielded 77 µg/ml in 1 ml of 100mM phosphate buffer pH 7.5. Protein concentration was determined by Nanodrop and the extinction coefficient was determined using the online tool expasy ProtParam using the molecular weight of 60.9 kDa. The concentration of the purified CHMO was low however it was only weakly expressed.

Although BVMO has been shown to have very broad substrate specificity, inhibitory concentration of the substrate as well as product inhibition has been well documented. Cyclohexanone has been reported to inhibit reaction rates at a concentration of between 0.2-0.4 g/L whilst caprolactone inhibits at much higher concentrations in the range of 4.5-5 g/L leading to activity reducing to zero (Alphand *et al.*, 2003). Assay results, although crude, appear in agreement with these reported values. At substrate concentrations of 0.5 and 5mM (0.49 g/L and 4.9 g/L respectively) the activity curve is largely the same. However at higher concentrations such as 50 mM and 500 mM (49 g/L and 490 g/L) the activity of the CHMO significantly reduces however does not disappear or diminish to a large degree despite such high substrate concentrations being used (Figure 4.11).

### 4.4.3 ε-caprolactone hydrolase (6-hexanolactone hydrolase) and 6-hydroxyhexanoate dehydrogenase activity

ε-caprolactone hydrolase (a.k.a 6-hexanolactone hydrolase, EC 3.1.1 - ) is the 3$^{rd}$ enzyme in the pathway and is responsible for the ring fission of the lactone into a hydroxyl acid by hydrolysis (Bennett *et al.*, 1988). Previous investigations into this enzyme have indicated it has high turnover numbers and is able to function over a broad range of pH values (Bennett *et al.*, 1988). Smaller chained hydroxy acids are known to spontaneously reconvert back into their respective lactone however it is more typical with shorter carbon chains and is not known to happen with 6-hydroxyhexanoate. Although other lactonases have shown rather broad substrate specificities, ε-caprolactone hydrolase was found to be very specific in that it has shown only minimal activity on several lactones other than ε-caprolactone and very high turnover on this alone. Since this reaction has no requirement for co-factors, absorbance assay's would be difficult and further investigation would likely involve GC or GC-MS and detection of 6-hydrohexanoate. However at this stage that would require

significant method development and so a simpler solution was to couple ε-caprolactone hydrolase with 6-hydroxyhexanoate dehydrogenase where only caprolactone and $NAD^+$ are provided and following the reaction based on the appearance of NADH thus confirming both enzymes are active. Very little work has been done on 6-hydroxyhexanoate dehydrogenase but it was noted by Donoghue *et al* that aside from it being labile in storage it exhibited slow enzyme kinetics (Donoghue and Trudgill, 1975). Both enzymes were again cloned into pET29a and His-tagged. Although both appeared to be highly expressed they purified out at low concentrations; ε-caprolactone hydrolase was at 207 μg/ml and 6-hydroxyhexanoate dehydrogenase was at 952 μg/ml. Again both co-eluted with a host cell protein. Never the less 20 μL of each enzyme was used in the reaction with the same range of substrate concentration. Results suggested a very high tolerance to substrate concentration with excellent activity seen even at 500mM caprolactone and a significant drop in activity at 50mM (Figure 4.12). No observable reaction was observed at 5mM and 0.5mM caprolactone. The cell lysate of BL21 (DE3) carrying empty pET29a showed no activity on ε-caprolactone therefore the activities observed are from the expressed enzymes.

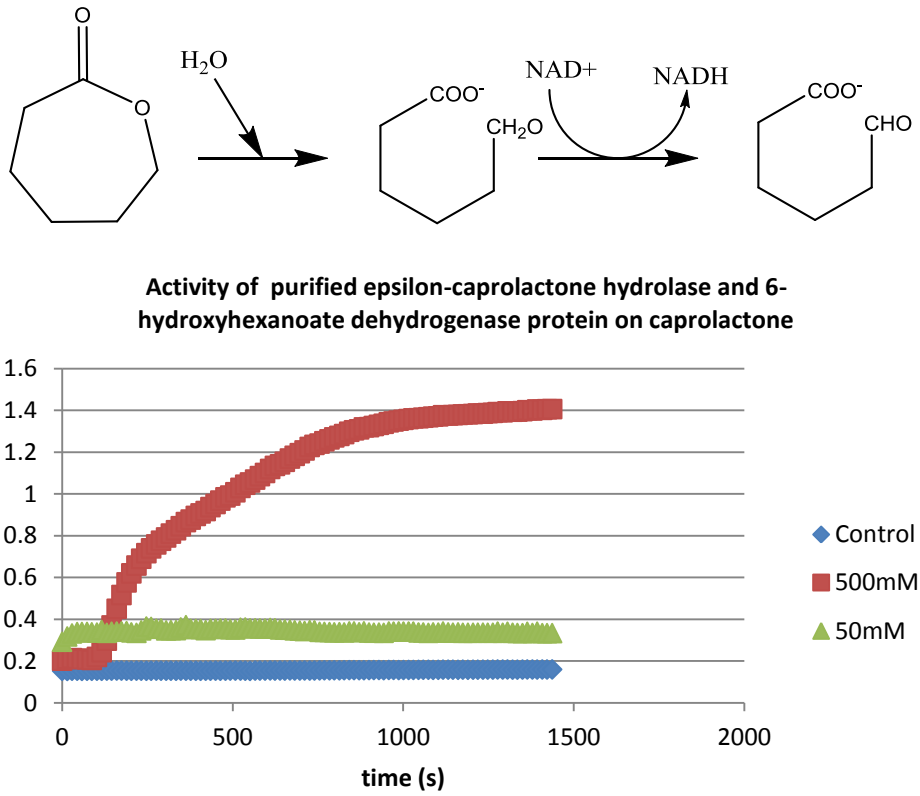



**Figure 4.12 Top is the schematic of the coupled reaction. Bottom is the absorbance assay of purified caprolactone hydrolase and 6-hydroxyhexanoate dehydrogenase on caprolactone in the presence of 0.5 mM NAD+ in 100 mM phosphate buffer at pH 7.5. The graph shows a large substrate tolerance and activity remains good at high substate concentrations. 50 mM caprolactone rapidly plateu's after a slight rise in acitivty. A schematic of the reaction is above the graph. Because the activity of the lactonase is difficult to detect by itself the reaction was coupled with 6-hydroxyhexanoate dehydrogenase. The lactone was supplied as the only substrate whilst any absorbance change is as a result of the product from the lactonase, 6-hydroxyhexanoate, being converted to 6-oxohexanoate and the conversion of NAD+ to NADH. There was no direct activity measured. Whether the enzyme was functional and how it behaved under different substrate concentrations was investigated. Therefore no figures for specific activity are given.**

To possibly identify if the kinetics of 6-hydroxyhexanoate dehydrogenase produce a similar response, cell lysate from pQR1552 cells expressing this enzyme was assayed in an identical fashion. *E.coli* BL21 lysates were used as a cotrol in order to see if the host cell is able to utilise the hydroxy-acid. Results in Figure 4.13 suggest that the BL21 host cell enzymes cells do indeed possess some background activity on 6-hydroxyhexanoate. This is to be expected due to the very reactive nature of the aldehyde produced. It is also indicated that 6-hydroxyhexanoate dehydrogenase may have low activity. Although Figure 4.13 shows

only 500 mM 6-hydroxyhexanoate, it did produce the largest absorbance change when compared to other lower substrate concentrations (e.g. 50 mM). Both the control and enzyme performed poorly at 50 mM. This suggests the activity is overall a poor functioning enzyme and could indicate a potential bottleneck in the pathway. How it behaves *In vivo* could be markedly different however since the optimum conditions haven't been explored. Instead, it can be said however that it does definitely possess some activity for the substrate compared to the control.

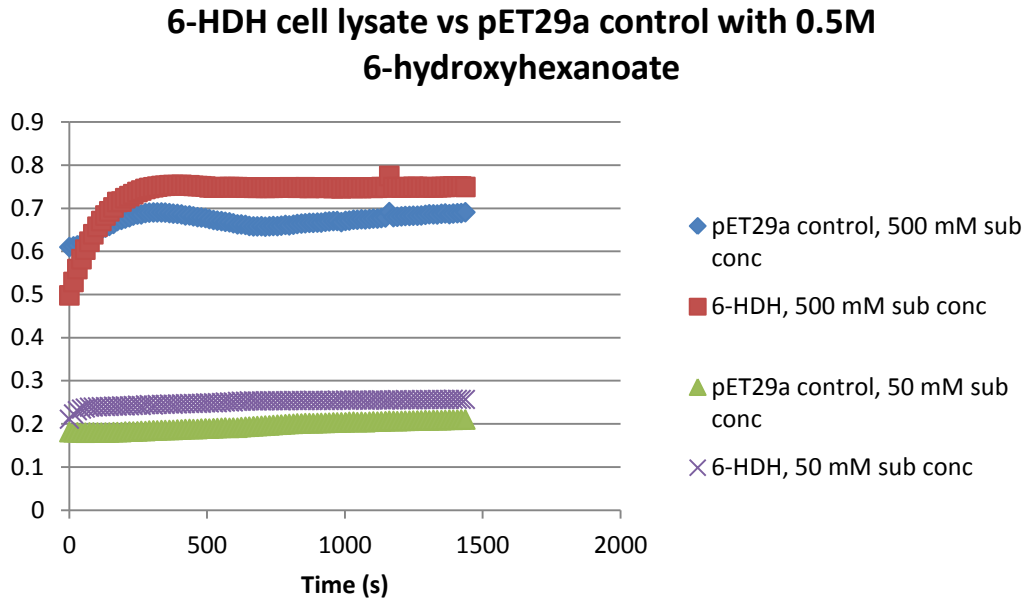



**Figure 4.13 6-HDH vs a pET29a cell lystate on 500mM and 50 mM 6-hydroxyhexanoate. At both substrate concentrations it can be seen that the 6-HDH enzyme performs marginally better than the *E.coli* BL29 (DE3) control however there does appear to be some background activity. The reaction schematic is seen as teh second half of the reaction in Figure 4.12. Any absorbance change was due to the formation of NADH and was detected via spectrophotometry. Activity wasn't measured, only an indication of activity was investigated i.e. is teh enzyme active and what are the crude substrate tolerances.**

# Chapter 5: ω-TAm cloning and screening

# 5 Results

## 5.1 Introduction

Following on from the successful operon cloning and activity indications the next phase of the project was to identify an appropriate Transaminase. Transaminases (aka Aminotransferase – E.C. 2.6.1 -) are of significant interest to industry because of their ability to produce optically pure chiral amines which are often used in the production of pharmaceuticals. Traditionally they have been produced using metals as well as having several reaction steps. Enzymatic routes however circumvent these issues. TAms are involved with the metabolism and generation of the 20 proteogenic amino acids and work by catalysing the reversible transfer of an amine group from an amino acid to a ketone or aldehyde (Wu *et al.*, 2017). The mechanism of action involves a redox reaction where by oxidative deamination of the donor is in conjunction with the reductive amination of the acceptor. PLP (pyridoxal 5' phosphate) is used as a prosthetic group; the biologically active form of vitamin-B6 and is regenerated after each reaction. The enzyme catalysis works by a ping-pong bi-bi mechanism and can be divided into two halves (Wu *et al.*, 2017). Firstly the lysine at the active site of the TAm forms a Schiff base with PLP forming an internal aldimine. After contact with the amine donor the bond between the cofactor and the enzyme breaks and the PLP forms a new Schiff base with the amine donor. Following a proton abstraction step, the amine group from the donor is transferred by hydrolysis to the PLP forming enzyme bound PMP (pyridoxamine 5'-phosphate) and a corresponding ketone. The second half of the reaction now proceeds in the opposite direction starting with the acceptor ketone/aldehyde being transferred the amine group from the PMP forming an amine product and regenerating PLP in the process (Sayer *et al.*, 2014).

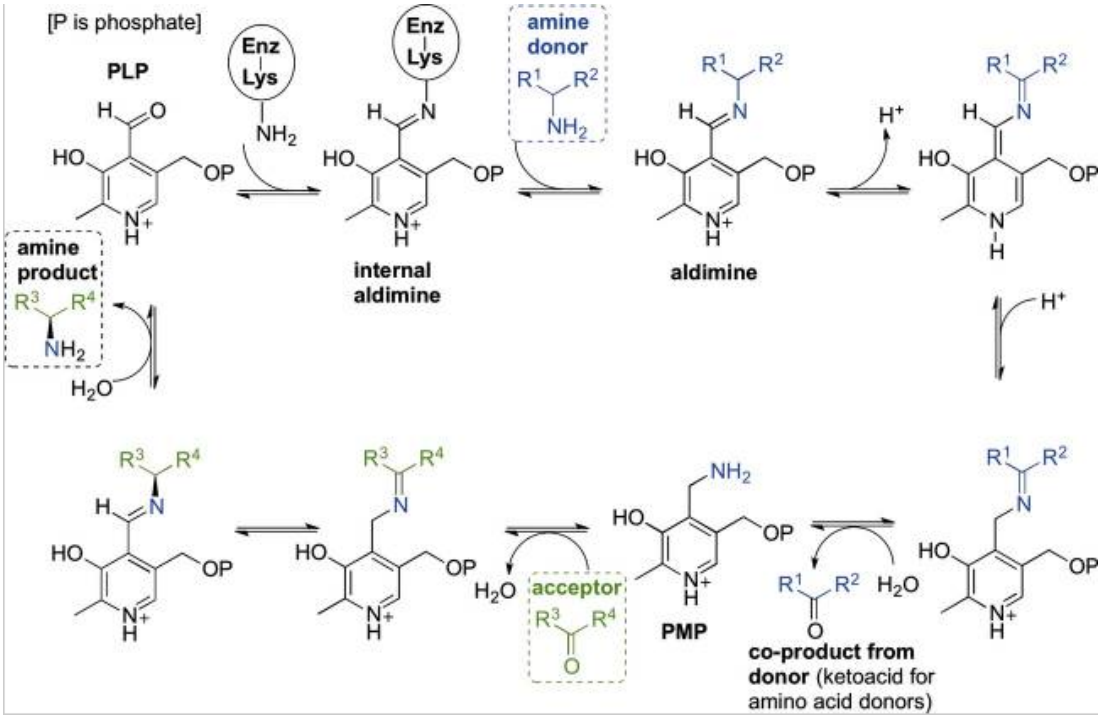



**Figure 5.1 Schematic of the transaminase reaction. Image copied from Sayar *et al.*, 2014.**

TAms are part of a family of PLP-dependent enzymes which are divided into groups depending on their fold configuration. TAms belong to family fold groups I-IV. Further investigations into sequence alignments led to the development of 6 sub-classes. Classes I, II, III and V are all the same fold type where I and II are aspartate and aromatic TAm's, III are ω-TAms and V are phosphoserine TAms. Another, broader classification system exists based on the position of the amine group on the carbon chain; α-TAms being able to react with amino groups on the α-carbon whilst ω-TAms act on distal amino groups. All TAms acting on β-amino acids are considered to be ω-TAms (Rudat, Brucher and Syldatk, 2012).

TAms enzymes typically have several limitations for their immediate applicability in an industrial setting. Firstly the reaction equilibrium tends to either favour the reverse direction (i.e. substrate direction) and so product yield can sometimes be low or equilibirum can be split in a ratio of 50:50 product to substrate. It is theoretically possible to push the reaction through by excessively loading the amine donor in the reaction to

push product formation, however, this can be limited by initial reaction rates. The second problem with TAms is that they can possess both substrate and product inhibition. There are several solutions to these problems including co-product removal using bi-phasic systems where an organic solvent is used to remove the product (Heintz *et al.*, 2017) or in the case of isopropanol as a donor with the co-product being acetone, an evaporation step could be included. Protein engineering can also be used to effectively decrease or remove the substrate or product inhibition by rational or random mutagenesis and high throughput screening (Park and Shin, 2013).

There are two potential targets in this pathway for a TAm to react. Upstream there is the ketone cyclohexanone produced from cylohexanol as well as 6-oxohexanoate, an aldehyde produced from 6-hydroxyhexanoate. In order to minimise any side product formation i.e. the transamination of cyclohexanone to cyclohexylamine, a suitable ω-TAm would need to be found with high activity on 6-oxohexanoate and minimal activity on cyclohexanone.

## 5.2 Screening 28 ω-TAm's for preferential activity on 6-oxohexanoate over cyclohexanone

### 5.2.1 Obtaining 6-oxohexanoate

6-oxohexanoate is a linear aliphatic aldehyde, its commercial availability is therefore limited because of its instability for long term storage. Therefore making it synthetically drives up the cost and consequently the price from retailers but also makes validation of experiments difficult since there will likely be a reducing amount of the aldehyde over time. Aldehydes typically are very reactive because of the large electronegativity of the oxygen atom and the presence of only one electron donating R group as opposed to two with a ketone. The large $\delta^+$ on the carbonyl carbon is subject to nucleophilic attack by even the

weakest nucleophile. Consequently even moisture in the air can act as a nucleophile leading to the rapid formation of either a carboxylic acid, its equivalent salt (adipic acid in the case of 6-oxohexanoate) or a gem-diol (Hydrate) in a nucleophilic addition reaction is possible. This posed a problem for performing screening studies to find the best Transaminase. However the methyl ester of 6-oxohexanoate is commercially available. Although the methyl ester group is at the opposite end of the molecule, it could be thought not likely to affect the screening in a large way and should provide a suitable alternative. However because this is a best available model there may be differences in the reactivity of the TAm's on the methyl ester compared to the *in vivo* 6-oxohexanoate. The degree to which this is so can be explored by chemically removing the methyl group leaving 6-oxohexanoate. This was done by acid hydrolysis (see Material and Methods section 2.5.2) and the structure confirmed by NMR (NMR work done by Damien Baud). The methyl ester is an expensive molecule to purchase and so an initial activity screen was done and compared to the chemically synthesised 6-oxohexanoate. A third method of making fresh 6-oxohexanoate was explored. Since the caprolactone hydrolase and the 6-hydroxyhexanoate dehydrogrenase were both cloned individually and expressing to a large degree these could be used to synthetically produce *in vitro* 6-oxohexanoate from caprolactone. Following Ni-affinity purification of each enzyme from induced cultures of both clones, they were used in equal volume of 100 µL each in a 5.3ml reaction and supplying a final concentration of 0.625 mM NAD, 1.11 mM PLP and 111 mM L-Alanine and 45 mM caprolactone at room temperature (see Materials and Methods 2.5.2). A sample was taken to follow the reaction using spectrophotometry and when it had completed the reaction mix was stored on ice. PLP and L-Alanine were included in the reaction at a higher concentration than was required in the TAm screen (section 2.5.3) because, although they served no purpose in the formation of 6-oxohexanoate, the reaction mixture would serve

as the large component in the Tam screen i.e. it contains buffer, PLP and the amino acid at the required concentrations once diluted by addition of TAm cell lysate.

### 5.2.2 Screening rational

Following the generation of 6-oxohexanoate, a screening assay was developed based on an established method of TAm screening previously explored in our group and applied to both cyclohexanone and 6-oxohexanoate screens. This method involved using S-Methyl benzylamine (*S*-MBA) as an amine donor and detecting the appearance of the acetophenone co-product via HPLC at a wavelength of 254 nm (Richter *et al.*, 2015). However since the most likely amine donor *in vivo* is an amino acid, Alanine was chosen for a number of reasons. Firstly in most microorganisms, it is most commonly only synthesised *in vivo* for the biosynthesis of proteins and is part of central metabolism and therefore likely to be one of the most abundant amino acids. Secondly its synthesis is very simple and requires only a glutamate-pyruvate TAm using pyruvate as a substrate which is provided by glucose metabolism (Akita, Nakashima and Hoshino, 2016). Consequently the assay design is simple and requires only the detection of pyruvate as the co-product after transamination. It is simple to detect the presence of pyruvate by the addition of a lactate dehydrogenase which is added to the reaction post transamination. No additional co-factor needed to be added in the 6-oxohexanoate screen since the NADH produced during the production of 6-oxohexanoate from 6-hydroxyhexanoate (Figure 5.2) however it is supplied in the cyclohexanone screen. The conversion of pyruvate to lactate is detected again by following it on a spectrophotometer. The decrease in absorbance as NADH is oxidised to $NAD^+$ is indicative of LDH activity. The more pyruvate present, and therefore the better the Transaminase on 6-oxohexanoate or cyclohexanone, the greater the degree of absorbance change.

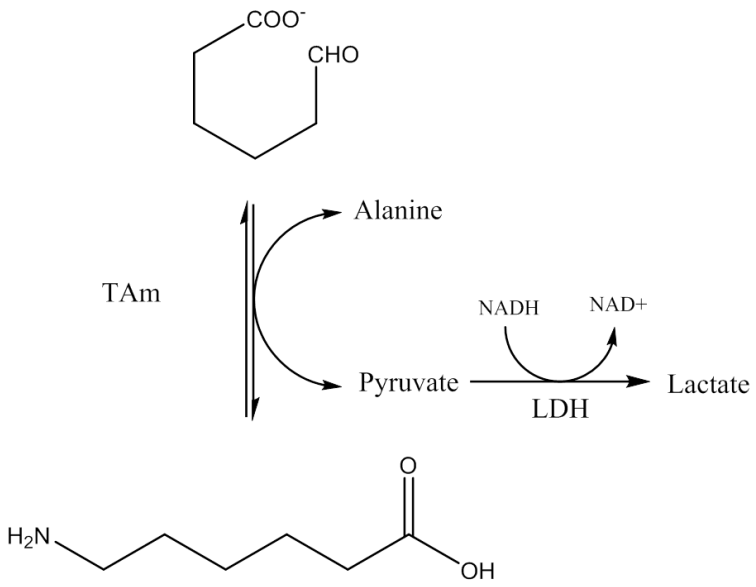



Figure 5.2 Schematic of the TAm screen design showing co-factor and enzyme pathway

### 5.2.3 Difference between Methyl-6-oxohexanoate and enzymatically produced 6-oxohexanoate

Since the model methyl ester substrate was available it was useful to see if there was an appreciable degree of difference between it and the biocatalytically produced 6-oxohexnaoate. This would either validate using the methyl ester as a viable substitute or not. It is to note that when using the methyl ester the absorbance decreased as expected because the LDH uses the NADH and oxidises it to $NAD^+$. Conversely however in the biocatalytically produced 6-oxohexanoate reaction, the absorbance actually goes up over time. The reason for this is that it is thought the excess (45 mM) ε-caprolactone that was used to generate 6-oxohexanoate before the TAm screen could only have used a maximum of 0.625 mM NADH to produce a stoichiometric amount of 6-oxohexanoate. When the Pyruvate from the TAm reaction is converted to lactate in the LDH assay, the $NAD^+$ released from this can then feed back into 6-hydroxyhexanoate dehydrogenase to further produce

6-oxohexanoate and NADH again (Figure 5.3). When compared to all other reactions, the TAm reaction is likely to be one of the slowest relative to the other reactions in the pathway. This is suggested based on previous research by Richter *et al* (Richter *et al.*, 2015) which showed that over the first 2 hours of a TAm reaction, conversion was low, ranging from 0-31 % conversion with most at the lower end of this range for non-natural substrates. Although this will differ with different substrate and enzyme combinations. This means that overall the LDH reacts quickly with the pyruvate producing $NAD^+$ which is subsequently used by 6-hydroxyhexanoate dehydrogenase (still present in excess) to produce 6-oxohexanoate which as shown previously (section 3.4.3) is, generally, a rapid reaction. However the TAm is unlikely to react with very much of the newly synthesised 6-oxohexanoate over the 40 minute spectral reading therefore producing minimal amount of pyruvate again in that time. Thus the absorbance decreases initially but it very quickly increases because of the generation of NADH which then presumably accumulates leading to an increased overall absorbance reading. This claim is substantiated by looking at an example absorbance reading of a reaction with the TAm expressed by pQR987 on 6-oxohexanoate (Figure 5.4). The first 200 seconds of the reaction shows a moderate decline in absorbance followed by a subsequent increase in absorbance and minimal decrease. It is likely that the initial decline is larger than is shown because of the time it takes to manually pipette before measuring.

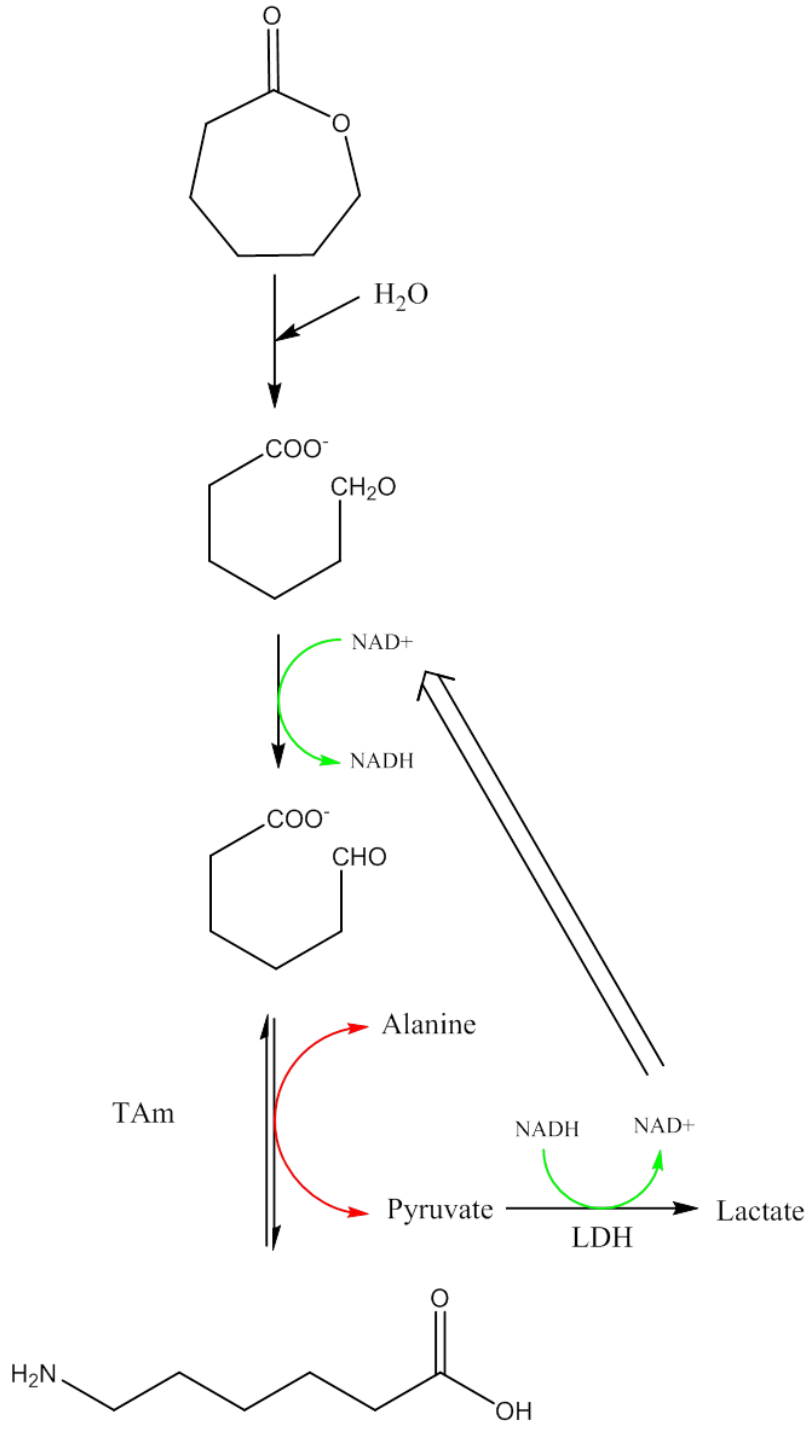



**Figure 5.3 TAm reaction schematic showing the fast reactions (green) and slower reactions (red).**

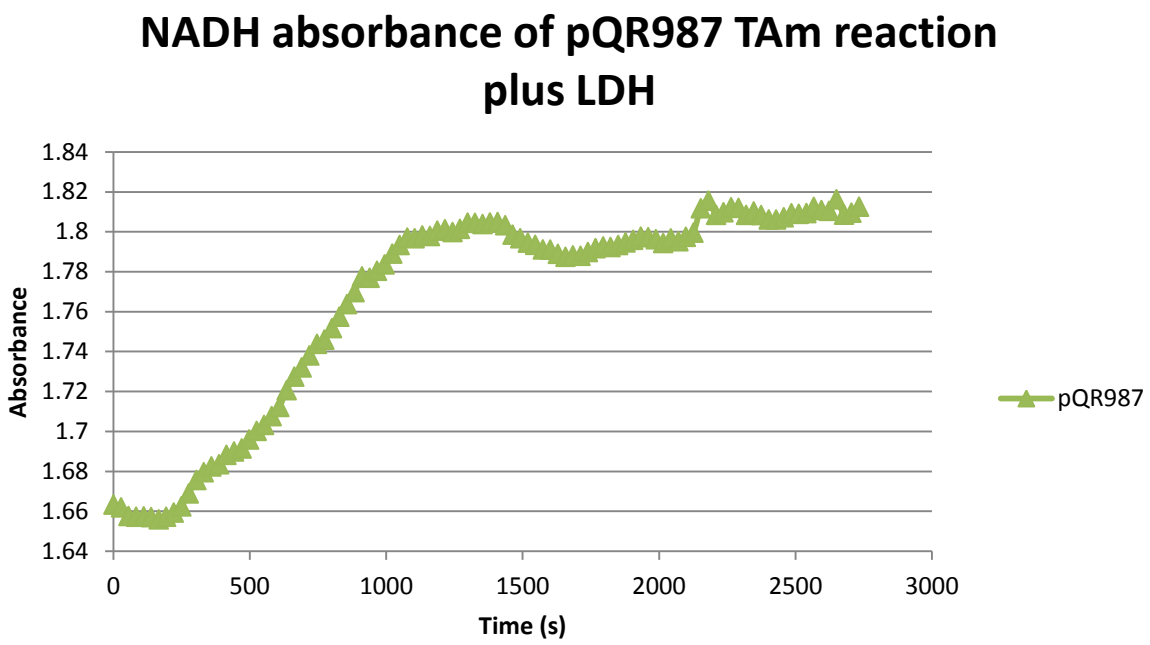



**Figure 5.4 An example absorbance spectra at 340 nm following the LDH assay of pQR987 TAm (*Bacillus licheniformis*) releasing pyruvate from alanine. Over approximately the first 200 seconds an initial decrease in absorbance due to the reduction of NADH can be seen which is followed by the increased absorbance as a consequence of the oxidation of the $NAD^+$ by in the dehydrogenase reaction of 6-hydrohexanoic acid to 6-oxohexanoic acid.**

The TAm clones were ranked with the largest absorbance change in the LDH assay being classified the highest ranked. Table 5.1 shows the results comparing the methyl ester to the biocatalytically produced 6-oxohexanoate. It becomes clear that although there are a few enzymes which appear to react in similar ways by having appreciable levels of closeness between the two ranks (e.g. pQR426 and pQR965), a considerable number showed opposing degrees of activity (e.g. pQR987). This cast doubt over the validity of the model substrate and therefore it would not be used in future experiments.

| Construct (pQR #) | Methyl-6-oxohexanoate (ranking of best performing TAm) | 6-oxohexanoate (ranking of best performing TAm) |
|---|---|---|
| 426 | 19 | 18 |
| 427 | 15 | 19 |
| 803 | 17 | 9 |
| 804 | 24 | 22 |
| 805 | 22 | 7 |
| 808 | 4 | 6 |
| 809 | 5 | 13 |
| 810 | 1 | 21 |
| 811 | 18 | 1 |
| 812 | 13 | 15 |
| 813 | 3 | 8 |
| 815 | 21 | 10 |
| CV2025 | 10 | 20 |
| 965 | 14 | 12 |
| 966 | 20 | 11 |
| 967 | 8 | 17 |
| 972 | 9 | 4 |
| 974 | 7 | 23 |
| 979 | 11 | 14 |
| 983 | 2 | 16 |
| 985 | 6 | 3 |
| 986 | 12 | 5 |
| 987 | 16 | 2 |

**Table 5.1 Numerical ranking of the methyl ester *vs* biocatalytically produced 6-oxohexanoate showing the differing results between the two. This could indicate that the methylated substrate is not a suitable alternative even though the methyl group is the opposite side of the molecule.**

### 5.2.4 Overall rankings

In order to identify the best candidate TAms a ranking system was designed to list the most ideal at postition 1 and least in the last position. Each experiment was done in triplicate on 6-oxohexanoate (either chemically or biocatalytically synthesised) and were ranked with the highest score being that with the largest average change in absorbance i.e. the largest absorbance change is ranked at position 1. In contrast to this, when screened on cyclohexanone, the Tams were ranked with the highest being minimal change from the start of the reaction and therefore showing no or minimal activity on cyclohexanone i.e. the smallest change from zero is ranked at position 1. The first experiment was the only one using biocatalytically produced 6-oxohexanoate and experiments 2 and 3 used the chemically synthesised version. This was due to the logistics of the arrival of the chemically synthesised methyl-6 oxohexanoate and the time to produce it and validate it. Table 5.2 shows the ranking tables and although there are certain experiments where values disagree with the others, overall there is no clear distinction between either method of obtaining 6-oxohexanoate i.e. broadly speaking, the biocatalytic experimental values do not appear to be significantly different to that of the chemically synthesised.

The rankings from each of the triplicate experiments were averaged. In order not to let obvious outlying values impact the final rankings, any large disagreements were not included. For example pQR426; experiment 2 ranking was significantly different (10) to either 1 (26) or 3(21). The final overall ranking was determined by adding the averaged ranks from both substrate screens (i.e. how a clone performed on cyclohexanone and 6-oxohexanoate) and ranked in ascending order, 1st being the highest combined ranking score.

Because this screen is likely to be inaccurate to a degree, the top three performing TAm's which exhibited some of the lowest activities on cyclohexanone whilst having some of the

highest activities on 6-oxohexanoate were selected. It is likely that one of the three would be a top performer *in vivo*. Also to note is that there was no weighting placed on either requirement i.e. a high activity on 6-oxohexanoate wouldn't substitute for having activity on cyclohexanone. For example pQR804 (4) was shown to have a slightly better activity on 6-oxohexanoate than pQR810 (1) however it was shown to be only 8th best for minimal activity on cyclohexanone. Despite being ranked 9th for activity on cyclohexanone, because pQR983 shows the highest activity on 6-oxohexanoate it ranked number 2.

Figure 5.5 is a graphical representation of the average absorbance change of each clone. The screens identified three potential candidates which would be suitable for cloning into both operon versions (native and non-native RBS) for metabolite characterisation. These were pQR810 (*Pseudomonas putida* KT2440, PP5182, 49.87 kDa), pQR983 (*Deinococcus geothermalis,* Dgeo_2743, 37.32 kDA) and pQR1014 (*Micrococcus luteus*, MLUT_00920, 50.2 kDa).

| | 6-oxohexanoate | | | | cyclohexanone | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| PQR # | Exp1 | exp2 | exp3 | Averaged rank | Exp1 | exp2 | exp3 | Average rank | Combined score | Final rank |
| pQR426 | 26 | 10 | 21 | 23.50 | 5 | 10 | 1 | 5.33 | 28.83 | 14 |
| pQR427 | 27 | 27 | 22 | 25.33 | 14 | 5 | 2 | 7.00 | 32.33 | 19 |
| pQR803 | 14 | 28 | 23 | 25.50 | 20 | 14 | 10 | 14.67 | 40.17 | 26 |
| pQR804 | 4 | 4 | 12 | 6.67 | 8 | 21 | 8 | 8.00 | 14.67 | 4 |
| pQR805 | 11 | 23 | 7 | 9.00 | 16 | 22 | 26 | 21.33 | 30.33 | 16 |
| pQR808 | 15 | 2 | 3 | 2.50 | 27 | 26 | 22 | 25.00 | 27.50 | 9 |
| pQR809 | 3 | 3 | 6 | 4.00 | 28 | 28 | 14 | 28.00 | 32.00 | 18 |
| pQR810 | 6 | 5 | 19 | 5.50 | 7 | 2 | 3 | 4.00 | 9.50 | 1 |
| pQR811 | 1 | 26 | 18 | 22.00 | 3 | 4 | 12 | 6.33 | 28.33 | 11 |
| pQR812 | 25 | 14 | 15 | 14.50 | 2 | 3 | 27 | 2.50 | 17.00 | 5 |
| pQR813 | 21 | 22 | 25 | 22.67 | 11 | 13 | 7 | 10.33 | 33.00 | 20 |
| pQR815 | 23 | 21 | 28 | 24.00 | 15 | 11 | 4 | 10.00 | 34.00 | 22 |
| CV2025 | 16 | 16 | 24 | 18.67 | 12 | 8 | 9 | 9.67 | 28.33 | 12 |
| pQR965 | 19 | 25 | 16 | 20.00 | 24 | 20 | 18 | 20.67 | 40.67 | 27 |
| pQR966 | 8 | 19 | 8 | 8.00 | 13 | 27 | 19 | 20.00 | 28.00 | 10 |
| pQR967 | 12 | 15 | 17 | 14.67 | 22 | 16 | 17 | 18.33 | 33.00 | 20 |
| pQR972 | 22 | 16 | 11 | 16.33 | 23 | 23 | 21 | 22.33 | 38.67 | 24 |
| pQR974 | 28 | 20 | 10 | 19.33 | 19 | 15 | 25 | 19.67 | 39.00 | 25 |
| pQR979 | 24 | 9 | 5 | 7.00 | 26 | 24 | 15 | 21.67 | 28.67 | 13 |
| pQR983 | 2 | 1 | 26 | 1.50 | 6 | 12 | 11 | 9.67 | 11.17 | 2 |
| pQR985 | 7 | 13 | 13 | 11.00 | 17 | 18 | 20 | 18.33 | 29.33 | 15 |
| pQR986 | 17 | 12 | 9 | 12.67 | 21 | 25 | 24 | 23.33 | 36.00 | 23 |
| pQR987 | 5 | 24 | 20 | 22.00 | 1 | 17 | 23 | 20.00 | 42.00 | 28 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| pQR1003 | 10 | 6 | 27 | 8.00 | 25 | 6 | 13 | 19.00 | 27.00 | 8 |
| pQR1010 | 18 | 11 | 4 | 7.50 | 9 | 19 | 28 | 23.50 | 31.00 | 17 |
| pQR1011 | 13 | 7 | 14 | 11.33 | 10 | 7 | 6 | 7.67 | 19.00 | 6 |
| pQR1014 | 9 | 18 | 2 | 5.50 | 4 | 9 | 5 | 6.00 | 11.50 | 3 |
| pQR1015 | 20 | 8 | 1 | 4.50 | 18 | 1 | 16 | 17.00 | 21.50 | 7 |

**Table 5.2 Numberical rankings of each TAm in triplicate, averaged and then the combined score was rankied with the lowest number being the best performing TAm. A rank of 1 in the 6-oxohexanoate screen means the most active whilst the rank of 1 in the cyclohexnaone screen means the least active. Yellow boxes indicate score which were not included in the averages due to extreme values. The combined score was then ranked itself so that the lowest score was ranked 1 i.e. the top performer. A list of each TAm encoded on the pQR contruct I.D's is available in Appendix B.**

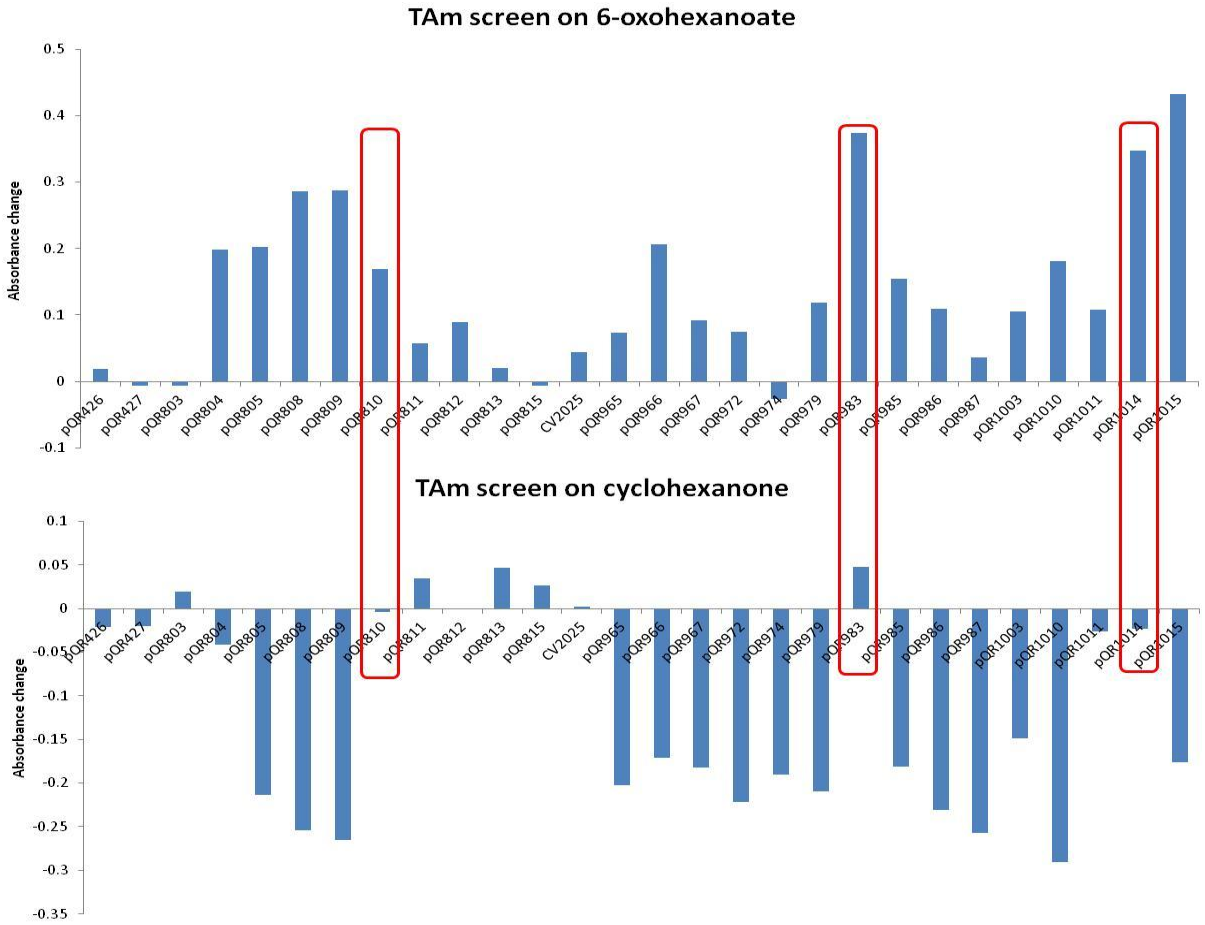



**Figure 5.5 Average absorbance change on each substrate. pQR810, pQR983 and pQR1014 are outlined as the top 3 performing candidates showing good indications of activity on 6-oxohexanoate and minimal activity on cyclohexanone.**

### 5.2.5 Cloning in PP5182, DGEO_2743 to operon A and B

The strategy for cloning the *Pseudomonas putida* and *Deinococcus geothermalis* TAm's into operon A and B was to first remove the *NotI* site at the end of *ChnD* so that the RBS ''AGGAG'' was immediately following it. This RBS was used for both operon versions since it is not from the *Acinetobacter* cluster and so by keeping the same RBS in both TAm, the idea that the native RBS have evolved to optimise flux through the pathway could still be tested. Overlaps complimentary for the TAm's were added to the primers whilst also linearising either plasmid. The observed yield of target PCR product was quite low and contributing factors as to why this could be are because of the length of primers and large template (Figure 5.6 and Figure 5.7). Also there was a prominent second and third band also produced. However after gel purification, enough was recovered to proceed with the CPEC reaction in a 1:2 molar ratio of insert (TAm) to vector (operon and plasmid genes). Both TAms' amplified with no co-product other than potential primer dimers and incorporated a *BamHI* restriction site flanking the TAm gene (creating the spacer betweek the RBS and the gene and also directly at the end of the TAm gene).

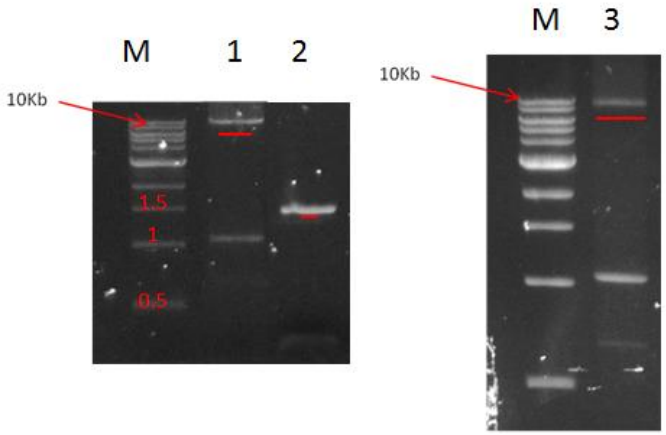



**Figure 5.6 Agarose gel of the inverse PCR amplification of the 9.2 Kb operon A (lane 1) and operon B (lane 3) and add homologous overlaps to for PP5182 and the PCR of a 1.36 Kb PP5182 *Pseudomonas putida* TAm (lane 2). The targeted fragment is underlined in red. M is a 1kb ladder (NEB,USA). In short the TAm in lane 2 will be CPEC cloned into linearised plasmids of operon A (1) and B (2).**

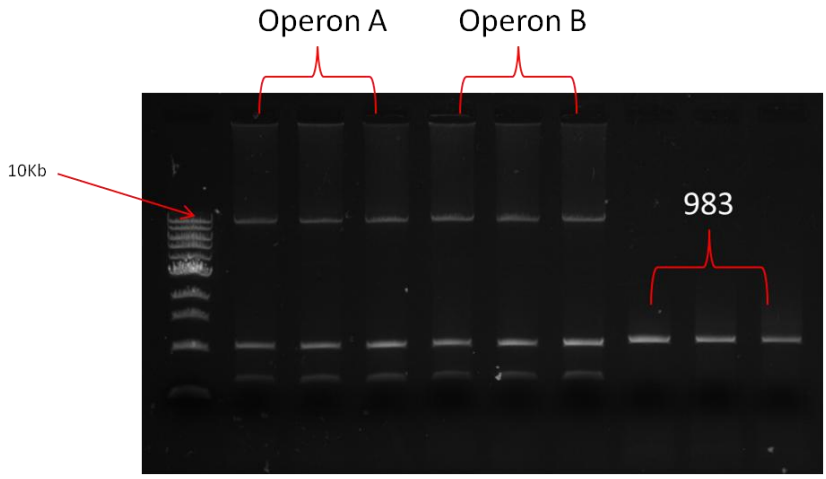



**Figure 5.7 Inverse PCR amplification of the 9.2 Kb operon A (pQR1543), operon B (pQR1548)plasmids. In addition a 1.02 Kb DGEO_2743 *Deinococcus geothermalis* TAm (pQR983) was amplified with overlaps for CPEC assembly. Note a significant side product of Operon A and B amplification.**

An interesting finding regarding CPEC assembly was that when inserting a single gene, the number of cycles required appeared to be important. Initially 26 cycles were used however zero colonies grew. However when the number of cycles was reduced to 15 at a cycle temperature of 55$^{o}$C, many colonies were observed. The reasons for this are unclear as to why with more fragments a higher cycle number appears to be important. It could be that the more cycles increased the formation of secondary structures which may result in degradation in the cell or it could be that the more cycles the higher the likelihood of binding to the *ori* or *Kan* genes resulting in a knock-out whereas with large constructs the chances are small because of the abundance of other separate sequences. The cycle number was not investigated further since it was only important that the assembly worked. *NdeI* and *BamHI* double digests were used to identify correctly assembled plasmids (Figure 5.8 and Figure 5.9). Figure 5.9 shows the digestion patterns of several miniprepped colonies of all 4 plasmid variants. From the assembly of *PP5182* minipreps 1a and 4b showed correct digestion and were transformed into BL21 (DE3). From the assembly of

*Dgeo_2743,* minipreps 1 and 4 were transformed into BL21 (DE3). The new clones were then renamed according to Table 5.3.

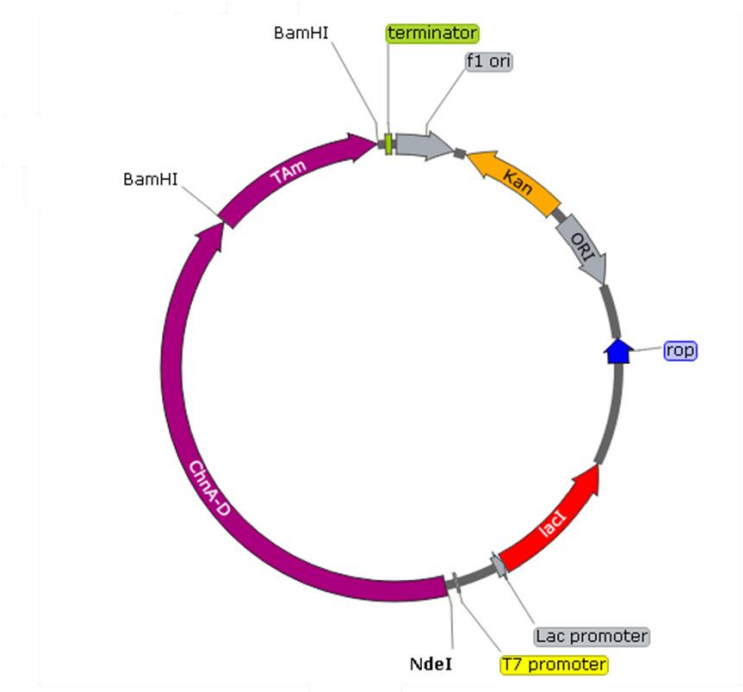



**Figure 5.8 A general plasmid map showing *NdeI* and *BamHI* sites, Both of which are used to confirm the correct assembly from CPEC. pQR1553, pQR1554, pQR1555 and pQR1556 were all designed to contain a *BamHI* site as a spacer between the RBS and the gene as well as at the end of the TAm. This was so it could easily be digested out in a single digest.**

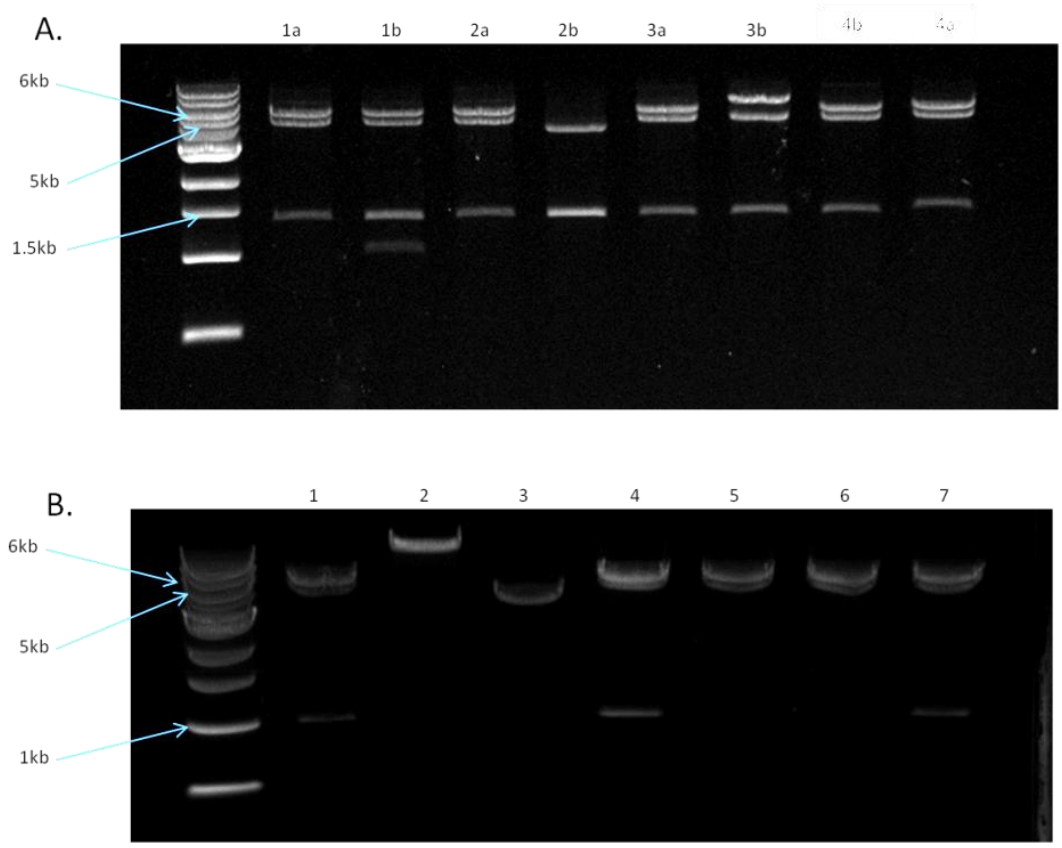



**Figure 5.9 *NdeI* and *BamHI* digested plasmid minipreps isolated from several colonies. Plasmids were double digested to identify which clones contain the operon, the backbone and the TAm. Top image (A) lanes are marked with a number to identify the miniprep and either the letter "a" or "b" refering to which operon version the TAm is in e.g. 3a is colony 3 with operon A assembly with *PP5182*. Clones 1a, 2a, 3a, 4b and 4a look to have digested correctly. Expected bands of succesful clones are 4.4kb (operon A), 5.2kb (plasmid backbone) and a thrid band for the TAm. *PP5182* is 1.36 kb. In the below image, B., plasmids 1-3 are operon A and 4-7 are operon B for *Dgeo_2743* assemblies. (*Dgeo_2743* band expected at 1.02kb). Clones 1, 4 and 7 look to have digested correctly.**

| Construct I.D. | Coding genes |
|---|---|
| pQR1553 | *ChnA - D* (natural RBS) + *PP5182* |
| pQR1554 | *ChnA - D* (artificial RBS) + *PP5182* |
| pQR1555 | *ChnA - D* (natural RBS) + *DGEO_2743* |
| pQR1556 | *ChnA - D* (artificial RBS) + *DGEO_2743* |

**Table 5.3 A table listing plasmid construct I.D numbers along with the coding genes within them**

An SDS PAGE of pQR1553 and pQR1554 (Figure 5.10) confirmed the expression of all *ChnA-D* genes as well as an over-expressed protein at the expected size of 49.87 kDa corresponding to *PP5182*. It is interesting to see that cyclohexanol dehydrogenase (the lowest band) shows a reduced expression in pQR1553 and a higher expression of *PP5182* relative to the other proteins compared to pQR1554.

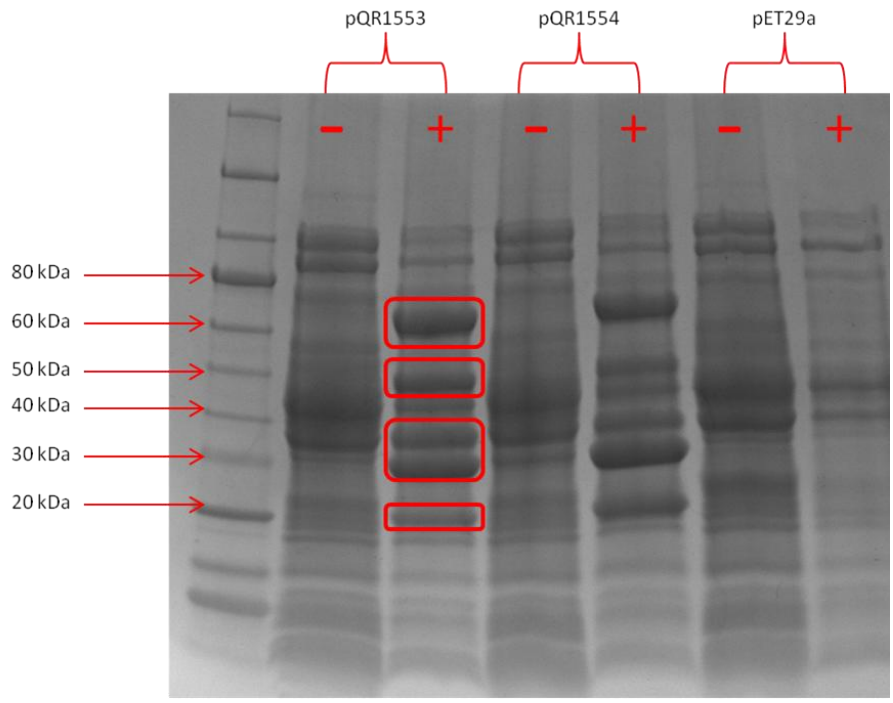



**Figure 5.10 SDS PAGE of pQR1553 and pQR1554. A 5th over-expressed protein band can be seen between showing the expected molecular weight of 49.87 kDa of PP5182 in both operon versions. A pET29a induced sample shows the presence of two highly expressing host proteins.**

An SDS PAGE of pQR1555 and pQR1556 (Figure 5.11) shows the successful expression of all 5 cloned enzymes under the control of a T7 promoter. Initial loading showed little resolution between *ChnD* and *Dgeo2743* because the difference in mass is only 0.1 kDa.

Upon loading of significantly less sample, two clear bands can be seen at expected molecular weights. Both clone variants showed excellent expression of all target proteins. The volumes loaded on a gel weren't normalised by loading a certain percentage of total protein and were loaded only on a fixed volume. Comparisons between clones are therefore difficult however relative expression conclusions can be drawn. In pQR1556, cyclohexanone monoxygenase and both 6-hydroxyhexanoate dehydrogenase and *Dgeo 2743* are arguably expressed to a lower extent than in pQR1555.

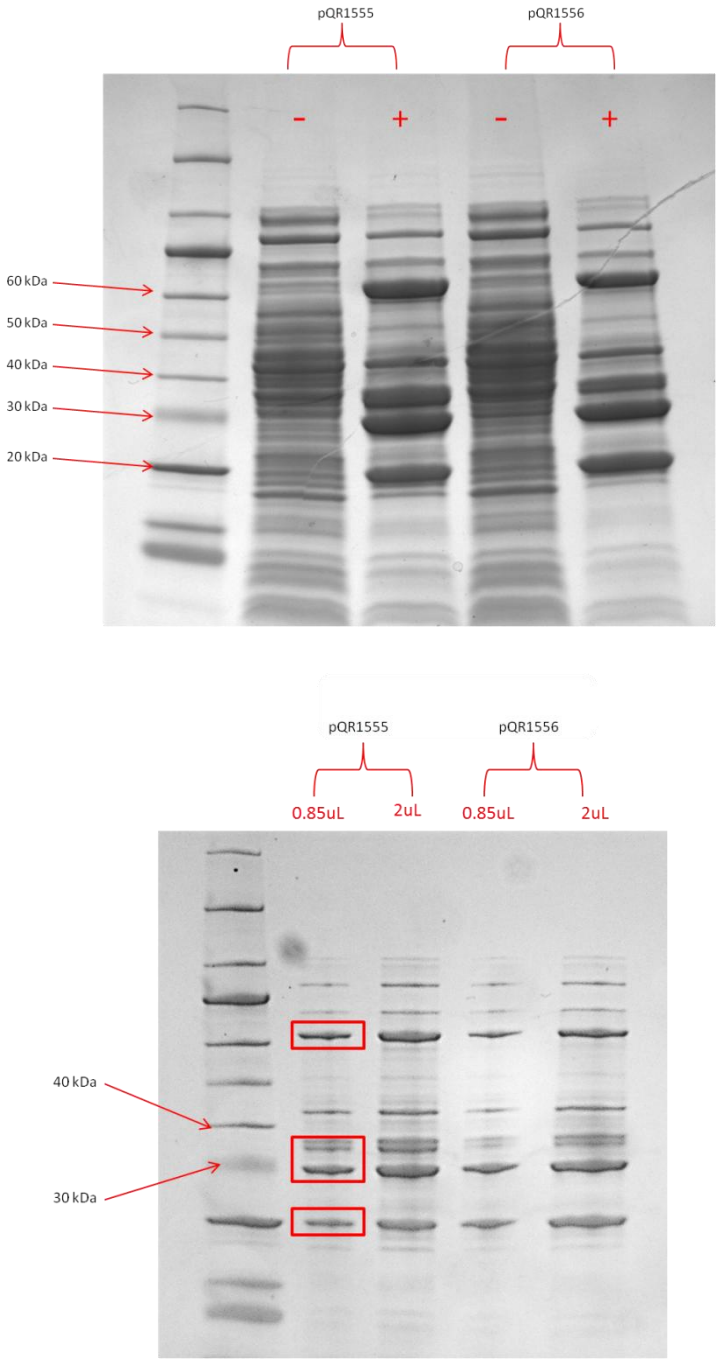



**Figure 5.11 SDS PAGE of pQR1555 and pQR1556. Top; non-induced and induced samples showing good expression of all cloned genes but with little resolution between Dgeo 2743(37.3 kDa) and ChnD(37.2 kDa). Bottom; SDS PAGE showing smaller sample loading where two bands with marginal size difference can be seen at less than 40 kDa. A prominent host cell protein above 40 kDa can also be seen.**

### 5.2.6 Cloning in *MLUT_00920* to operon A and B

Initially the strategy to clone *MLUT_00920* was to follow the same as with *Dgeo_2743* and *PP5182* however the PCR to linearise the vector and *MLUT_00920* whilst adding on overlaps failed repeatedly. Following this primers were designed to only linearise both parts. This was successful with *MLUT_00920* and a significant band of 1.4 kb was amplified and used as a template to add vector homologous overlaps to using the same primers as were previously used directly on the pQR1014 plasmid (Figure 5.12). This showed that there were problems when using long primers where only part of it bind initially to the circular template but could be circumvented by linearising the plasmid and then adding on overlaps. With regards to linearising the operons plasmid, only a very faint band was amplified of either operon A or B (Figure 5.13). After gel overloading with the sample and purifying the correct band, only very small quantities were recovered. The purified PCR product was then used for further amplifications as a template using the primers to add homologous overlaps for *MLUT_00902* however this was unsuccessful.

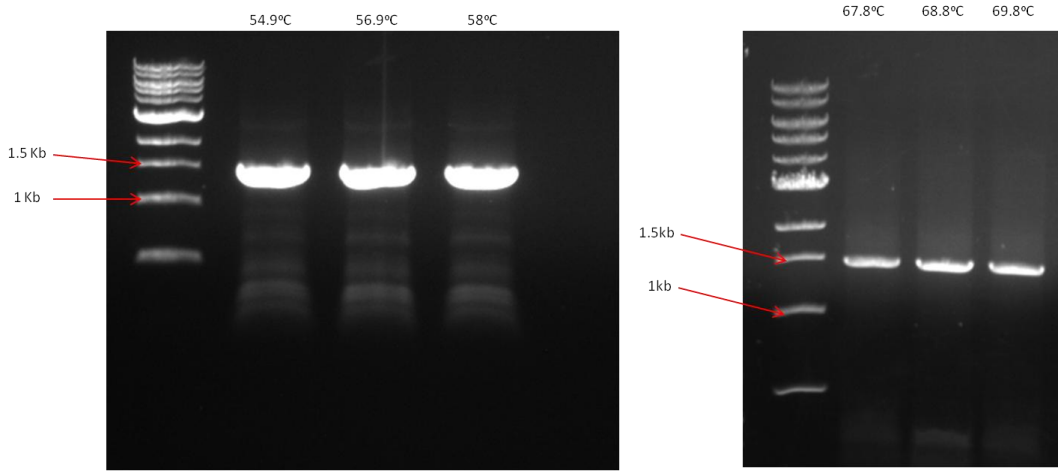



**Figure 5.12 . Left; linear amplification of MLUT_00902 with 3 different annealing temperatures showing a band of approximately 1.4 kb. Right; amplification using linear template to add vector overlaps**

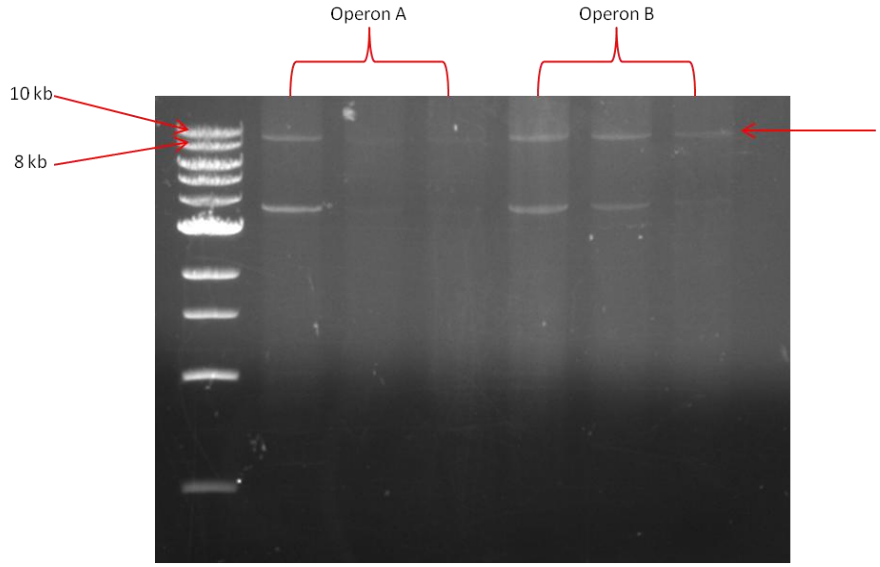



**Figure 5.13 Linearising amplification of operon A and B producing a faint band of 9.2 kb. Other off-target products were also produced. Further amplification using purified linear plasmids to add on *ChnD* overlaps failed.**

The result so far meant that in terms of a CPEC reaction, only the *MLUT_00902* gene contains homologous overlaps of 21bp at each end containing *BamHI* flanking sites for *ChnD* and the vector backbone. The operon plasmid was only linearised and amplifications to add overlaps were unsuccessful. A CPEC reaction was attempted with these two parts (Figure 5.14) with 10 and 15 cycles and 52°C and 55°C with a vector to insert ratio of 1:5.


**Figure 5.14 CPEC reaction attempted using only MLUT_00902 containing homologous overlaps to ChnD.**

Transformation of the CPEC reaction into cloning strain *E.coli* 10β cells produced plentiful colonies for operon A at 10 cycles annealing at 55°C and a small number of colonies with operon B assembly. Following overnight growth the plasmids from these colonies were miniprepped and digested with *NdeI* and *BamHI* for restriction mapping. Expected bands

were a 5.2 kb pET29a backbone, a 4.3 kb operon fragment and two fragments of *MLUT_00902* because *BamHI* cuts the 1.4 kb gene into a 0.8 kb and a 0.6 kb fragment (Figure 5.15).

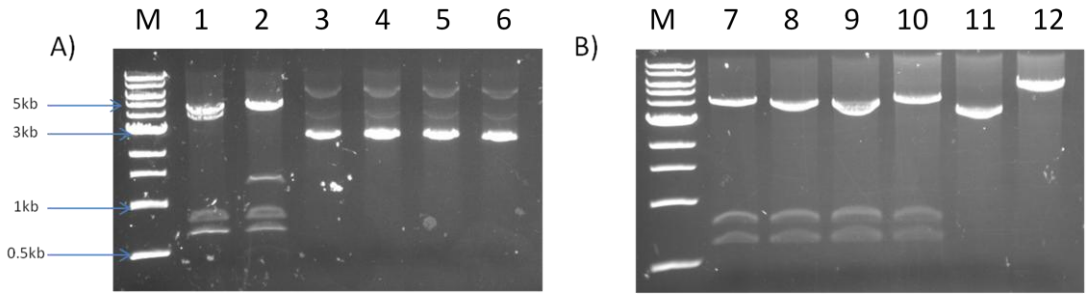



**Figure 5.15 A) plasmid digestions of six operon A plus MLUT_00902 assemblies miniprepped from individual colonies and digested with *NdeI* and *BamH I* (lanes 1-6). B) Six operon B plus MLUT_00902 assembly colony digestions with *NdeI* and *BamHI* (lanes 7-12).**

Although *MLUT_00902* was in 2/6 operon A colonies and 4/6 operon B colonies, only 1 operon A colony showed the presence of 4 bands, both of which appeared to be slightly lower than expected. No colonies of operon B showed correct digestion. Given that sometimes incorrect digestion or other factors e.g. protein binding, can affect the appearance of bands on a gel, colony 1 from operon A and colony 7 from operon B were transformed into *E.coli* BL21 (DE3) and expressed to identify if, given that *MLUT_00902* appears in both, it expresses along with the *Acinetobacter* genes. Figure 5.7 however shows that neither plasmid expressed any of the proteins and indeed looks to have removed *ChnA-D* genes from both plasmids.

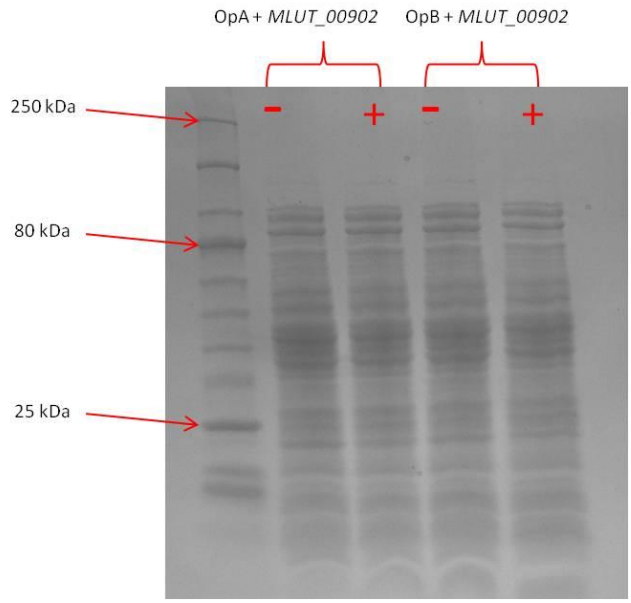



**Figure 5.16 E.coli BL21 (DE3) expression shows removal of *ChnA*-D genes and no expression of MLUT_00920.**

The reason for *MLUT_00920*'s, failure at CPEC cloning could be attributed to its high GC content (72%) compared to *Dgeo_2743* and *PP5182* being 64% and 62% respectively. The high GC content could have made the gene ''sticky'' and bind in non-specific regions. A change in cloning strategy was needed and it was decided that standard restriction cloning was the most appropriate method. However, more linear operon plasmid was needed than was currently able to be produced because not only was PCR amplification low but spin column purification also seemed to recover low single digit nanogram per microlitre concentrations. To that end the PCR amplification was investigated to optimise linear plasmid yield. An extended extension time of 50 s/kb was investigated following NEB's advice for amplifying longer fragments however this also failed. Long PCR primers were also tried with the reasoning that a 2 step PCR could be used and that the likelihood of large primers to bind to the plasmid was high also failed to produce any linear fragment. Fresh Q5 polymerase was also bought however the result was also the same. Finally a reduced extension time of 15 s/kb was tried and Phusion™ and Q5 polymerases were compared using previous primers which had failed with longer extension times.

Interestingly there were significantly different results between both polymerases with Q5 producing bands in two of the three annealing temperatures and Phusion™ producing zero bands at any of the same annealing temperatures (Figure 5.17). Also it is interesting that a shorter extension time of 15 s/kb (using the same primers and template as used at 50 s/kb) produced much better 9.6kb product which contradicts NEB's advice for long fragments.

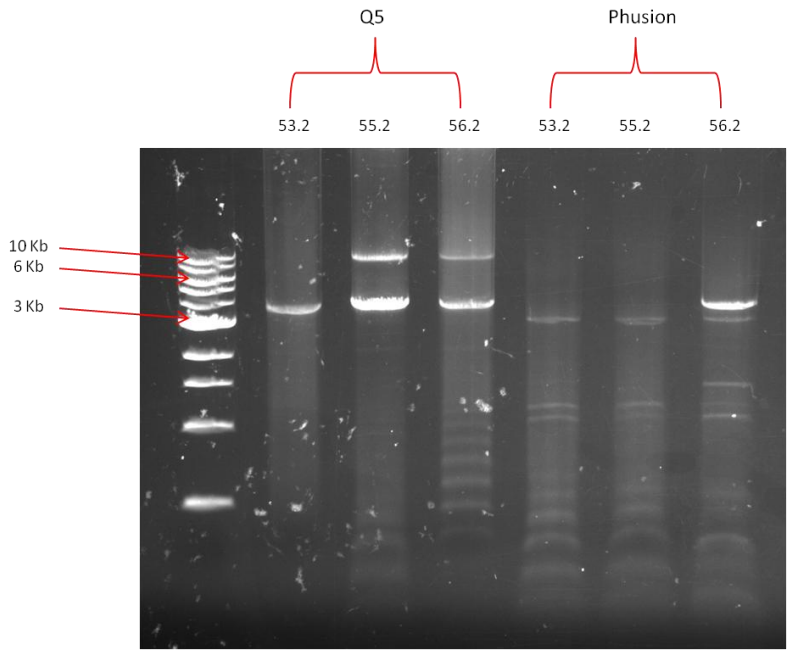



**Figure 5.17 PCR amplification comparing Q5 and Phusion™ polymerases at 15 s/kb with primers previously used at 30 s/kb extension time.**

In order to clone in *MLUT_00902* standard restriction cloning was the next cloning strategy. After identifying BbvcI and *AvrII* restriction sites as being candidates due to their absence from both vector and insert as well as reported 100% activity in the same buffer, primers were designed to linearise both operon A and *MLUT_00902* whilst adding *AvrII* and BbvcI restriction sites to the ends of each fragment (Figure 5.18).

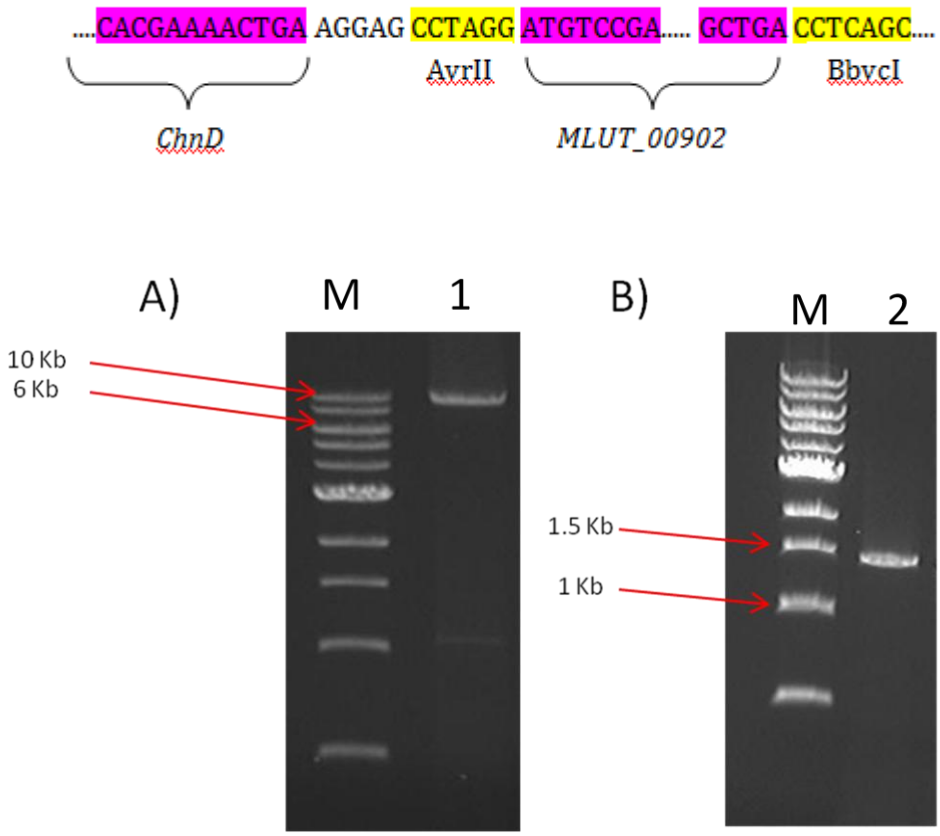



**Figure 5.18 A) PCR amplification of Operon A (lane 1), B) PCR amplification of MLUT_00902 (lane 2) adding in *AvrII* and BbvcI sites to both fragments. The marker (M) is a 1kb ladder from NEB, USA.**

After digesting and ligating each fragment, several cloning colonies were grown overnight and miniprepped. They were then triple digested with *NdeI*, *Bbvc*I and *AvrII* to check for correct assembly (Figure 5.3). All 8 colonies showed correct assembly of the plasmid and one was transformed into BL21 (DE3) for expression and renamed pQR1557.

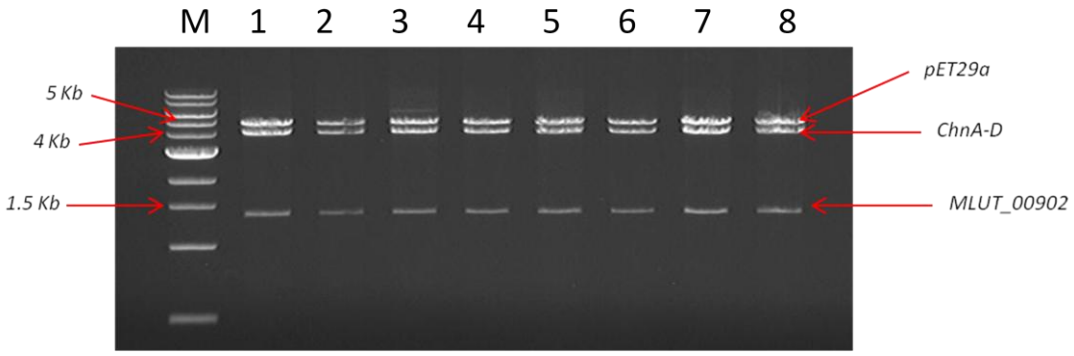



**Figure 5.19 *NdeI*, *BbvcI* and *AvrII* triple digest showing a 5.2 Kb fragment (pET29a), 4.4 Kb fragment (*ChnA*-D) and a 1.4 Kb fragment (MLUT_00920). All 8 colonies (lanes 1-8) show the correct digestion patterns expected. Marker lane is a 1 kb ladder (NEB, USA)**

Figure 5.20 shows the SDS PAGE of pQR1557. Five over-expressed protein bands can be seen however *MLUT_00920* looks to be under the 50.2 kDa predicted size. An SDS PAGE of pQR1014 shows an over-expressed protein band at the same sub-50 kDa size which is evidence that it has been successfully cloned into pQR1557. It could be that the protein hasn't properly denatured and may require more rigorous denaturing and so has travelled further in the gel than predicted. However since the original pQR1014 expression shows an identical band size then it appears to have been cloned successfully. *MLUT_00920* was not cloned into operon B because of poor analytical performance of pQR1557 (see next chapter) and would likely not be significantly different.

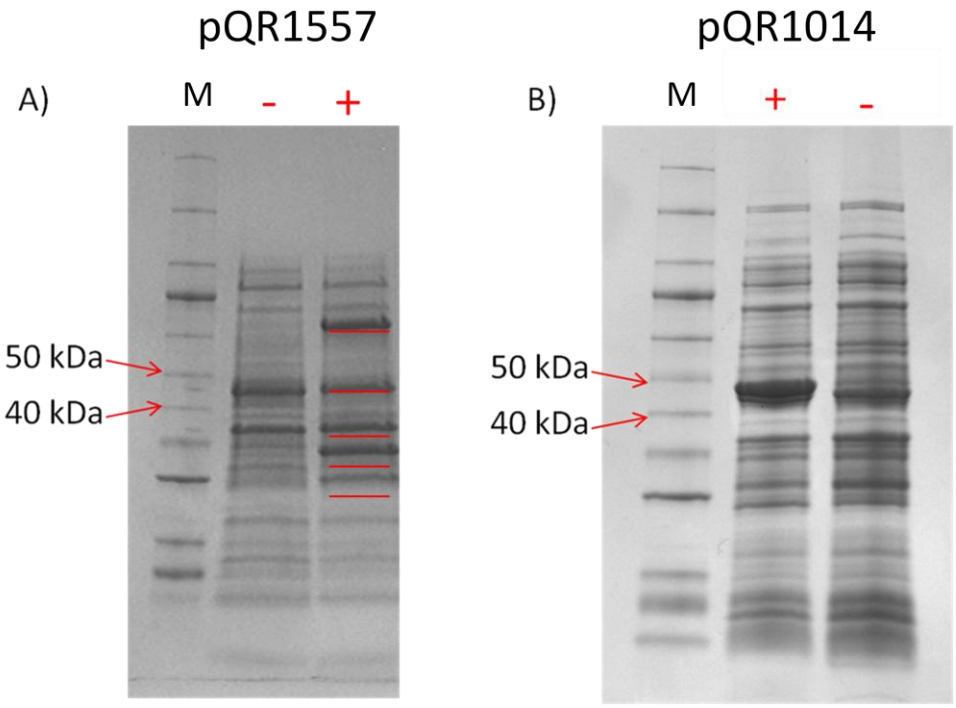



**Figure 5.20 A) an SDS PAGE of pQR1557. Non-induced (-) shows potentially leaky expression of ChnD however all 5 genes can be seen expressing (induced lane +) (underlined by a red line). B) pQR1014 SDS_PAGE showing a prominant band between 40 and 50 kDa. This band is identical in size between pQR1557 and pQR1014 showing successful expression of the TAm.**

# Chapter 6: Metabolic intermediate detection by GC-MS

# 6 Results

## 6.1 Introduction

Analytical detection of non-aromatic amines has been known to be challenging for several reasons. Firstly because no chemical groups are naturally present which are UV active, that absorb UV or emit fluorescent light, techniques such as HPLC with UV detection is generally not possible (Cramer, Verma and Kiselewsky, 2011). GC (gas chromatography) offers several advantages for analytical measurements. It works by increasing the temperature of the oven and as the analyte evaporates it is carried in a constant gas flow and interacts with the stationary phase of the GC column. Depending upon the degree of interaction the analytes will elute at different times and temperatures e.g. low volatile large compounds eluting later. It is inherently simple, has excellent resolution capabilities, is highly sensitive and has short run times. However free amines are basic molecules, and short chained primary amines being the most basic due to the larger influence of the amine group which subsequently form a large dipole molecule. These large dipole moments have a strong capacity to interact with acidic silanol groups and siloxane bridges found in GC columns. In addition free amines are likely to oxidise under high temperature conditions and could also cyclise. These effects are typically seen by peak tailing, low detector response and low reproducibility. In order to improve detection, amine compounds are often chemically derivatised to reduce polarity and increase volitility, sensitivity and separation. There are several methods with which to derivatise amines notably by acylation, permethylation and carbamate formation (Kataoka, 1996). One of the most popular is silylation. Silylation agents are typically BSA (N,O-Bis(trimethylsilyl)acetamide), MTBSTFA (N-methyl-N(*tert*-butyldimethylsilyl)triflouoacetamide) and BSTFA (N,O-bis(trimethylsilyl)acetamide). To compensate for low reactivity of silylating reagents to amines (order; Alcohol>phenol>carboxylic acid>amine>amide) TMS (Trimethylchlorosilane) is added as a

catalyst. BSTFA/TMCS reacts with active hydrogens and replaces them with a trimethyl silyl group (- $Si(CH_3)$) group which on amino acids or similar compounds (e.g. 6-ACA) is added onto both the carboxyl and the amine group (Figure 6.1). It is very reactive with moisture and water and so is only used in the absence of either.

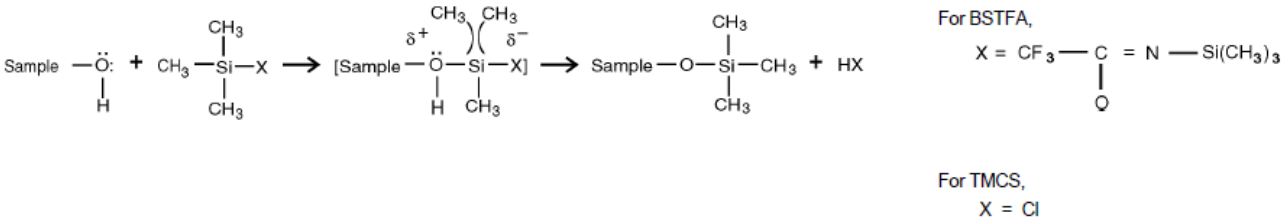
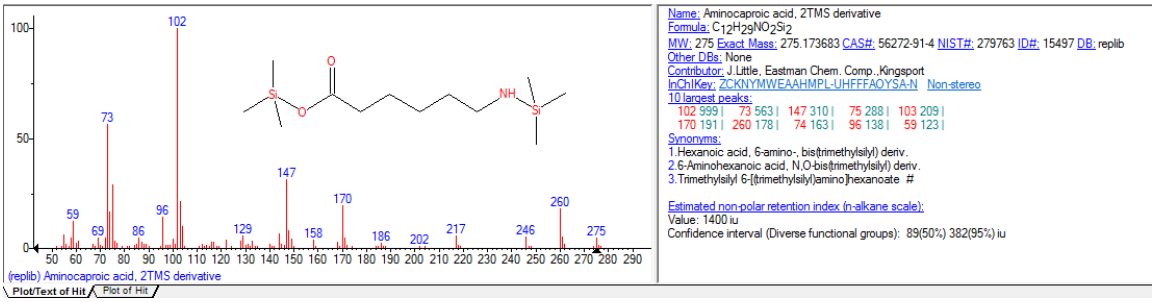



**Figure 6.1 Image copied from Sigma Aldrich's (BSTFA + TMS, 99:1) product specificiation. The sample with the reactive group to be derivatised is mixed with BSTFA with TMCS where the active hydrogen from the sample is replaced with the alkylsilyl group in a nucleophillic attack on the silicon atom forming a transition state. The leaving group then reacts with the active hydrogen from the sample by donating the trimethyl silyl group.**

## 6.2 Solvent selection, derivatisation and GC-MS intermediate analysis

Since the intermediate molecules of the pathway are polar, a suitable polar solvent was required which was compatible with GC-MS analysis. Although water is an excellent polar solvent, it's low volatility would cause condensation on the GC column at low temperatures and interfere with chemical-film interactions and possibly leave impurities e.g salts, metals. Cyclohexanol, cyclohexanone and ε-caprolactone all dissolved in a range of organic solvents and all up to 6-hydroxyhexanoic acid would readily dissolve in pyridine as well as methanol. Cyclohexylamine and 6-aminocaproic acid were much more problematic since no solvent (including harsh solvents such as DMSO) were found to dissolve the free compounds. The only solvent which would readily dissolve free 6-ACA was water. Ethyl acetate was chosen as a solvent which showed excellent ability to extract the first three intermediates from water into an hydrophobic layer and this was confirmed by GC. Ethyl acetate, although polar, lacks the ability to hydrogen bond and is therefore insoluble in water. 6-ACA was found to only dissolve in ethyl acetate when derivatised with excess BSTFA/TMS for 1 hour

at 70°C. When a solution of 6-ACA at 1 g/L was derivatised and injected on the GC-MS, a peak was confirmed by MS to be 2TMS derivatised, once on the amine and once at the carboxyl terminal (Figure 6.2). No derivitisation occurred at room temperature or when left




for extended periods of time e.g. 24-48 hours at room temperature.

**Figure 6.2 Spectra showing 2TMS derivative of 6-ACA detected at 8.1 minutes**

In conclusion the solvent of choice for the first three intermediates (cyclohexanol to ε-caprolactone) is to solvent extract into Ethyl Acetate. The rest of the intermediates will be derivatised in excess BSTFA and 1% TMCS at 70°C for 1 hour since it was found that once derivatised, the compounds then become soluble in Ethyl acetate (Figure 6.3)

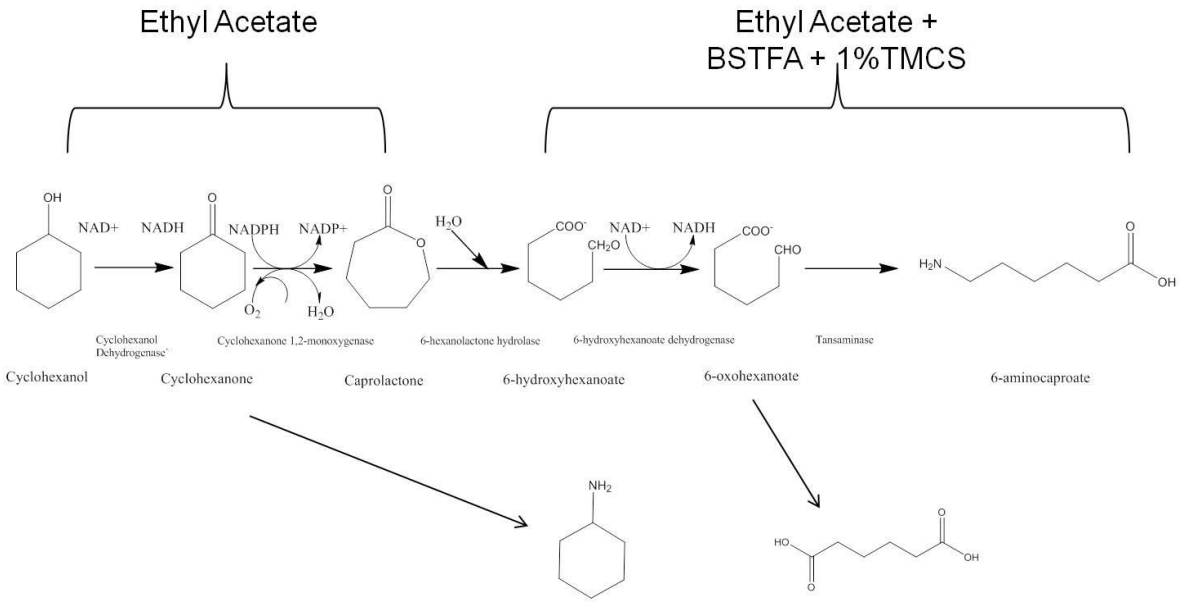



**Figure 6.3 Schematic of the solvent of choice and derivatisation reagent for identification of each intermediate. Cyclohexanol to caprolactone were solvent extracted into ethyl acetate whilst 6-hydroxyhexanoateto 6-aminocaproate including the side products of cyclohexylamine and adipic acid were derivatised with BSTF + 1% TMCS in ethyl acetate.**

### 6.2.1 Cell culture processing and preliminary result for *in vivo* 6-ACA detection

Figure 6.4 shows a schematic for the method of processing from a cell culture. Firstly after the cells have been inoculated, induced and spiked with cyclohexanol, samples were take the following day and initially processed immediately. Initially the aim was to find any evidence of 6-ACA production and so a culture of pQR1556 cells were grown, induced and 200 mM of cyclohexanol was added and left overnight. This concentration was chosen because, although it is very high, previous enzyme work (chapter 3.4) showed their capacity to catalyse high substrate concentrations. The cultures were then sonicated to lyse the cells, centrifuged and then processed (Figure 6.4). For this initial experiment the samples were dried in a drying oven overnight and then reconstituted in Ethyl Acetate and BSTFA ( and 1% TMCS) and then derivatised.

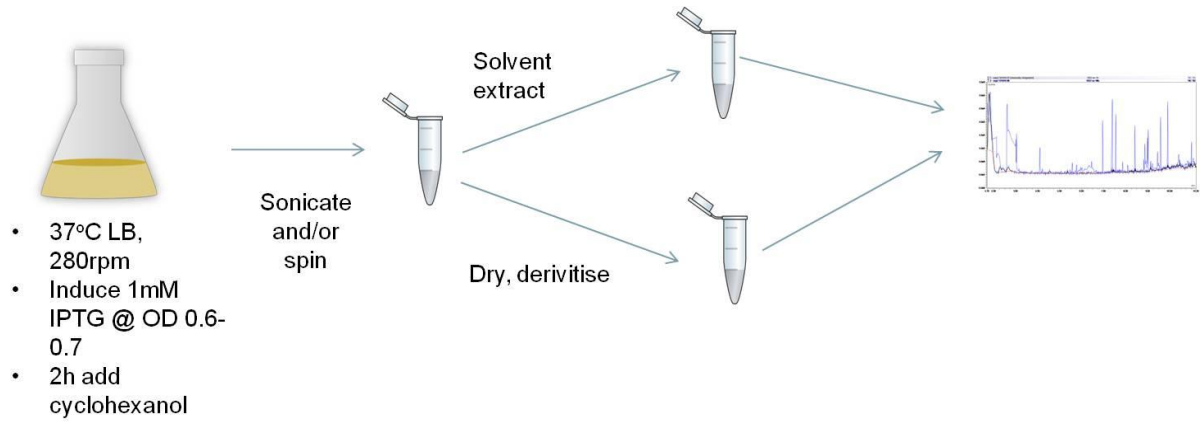
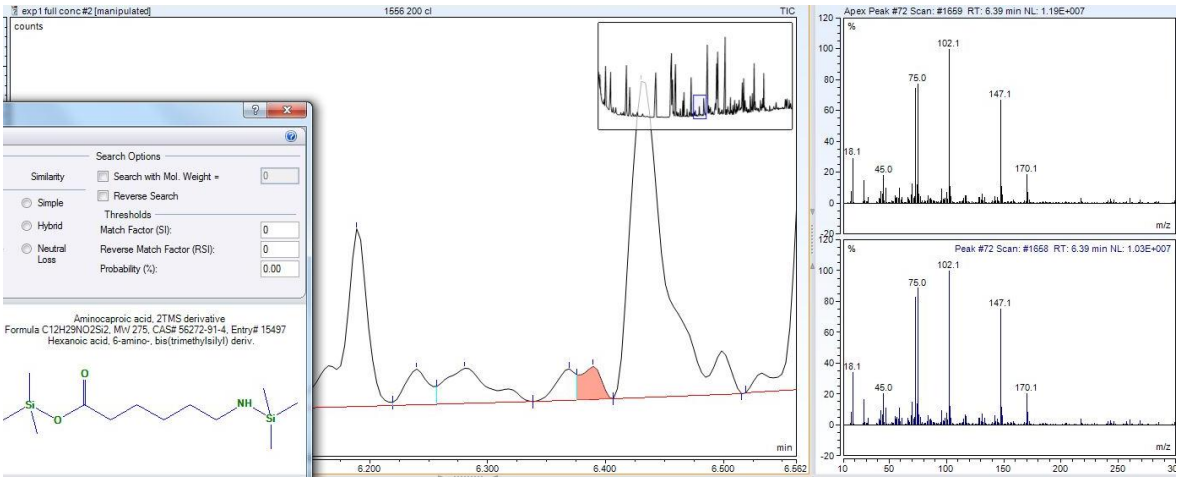



**Figure 6.4 Cell processing for amine detection. The figure shows the method of processing from cell culture samples to GC-MS ready samples. The first stage (either spin or sonicate) indicates which fraction of the broth is being analysed.**

At this point the GC method required for running cell culture samples needed to be investigated further since a standard won't show if a longer run with different carrier gas flow rates or what split ratio is required once the intermediate is present in a complex mixture so a best guess was used and various split ratios tested. Nevertheless a small peak at a split ratio of 5, confirmed by GC-MS as 2TMS 6-ACA, was identified at 6.4 minutes

(Figure 6.5). Notably there was still a significant cyclohexanol peak found which would suggest inadequate uptake by the cells and also incomplete evaporation at 37°C. It is reported that as temperature increases, cyclohexanol becomes more soluble in water and at 30°C it is soluble at 4.2g/100ml of water owing to the large non-polar portion of the molecule (PubChem Database; accessed 07/09/2017) 200 mM is equivalent to 2.003 g/100 ml and so should readily dissolve at 37°C. Compared to a control it appeared that the cyclohexanol also had an effect on the cell growth. However given that 6-ACA is identified, a GC-MS method could be improved to achieve better peaks as well as investigating other parameters to improve the final concentration.

**Figure 6.5 initial preliminary results gained from 200 mM cyclohexanol with pQR1556 cells identified a small peak which was confirmed as 6-ACA by MS.**




## 6.2.2 Standards preparation and the difficulty of 6-ACA

In order to quantify the intermediate peaks a standard curve of various incremental concentrations of a pure (>99% where could be bought, >95% for 6-hydroxyhexanoate) intermediate was used to generate a linear curve of concentration vs peak area. This was generated on the Chromeleon™ software. For intermediates from cyclohexanol to ε-caprolactone a 10 g/L stock concentration was made in ethyl acetate and then dilutions made into ethyl acetate were made ranging from 20 mg/L to 120 mg/L and it generated

linear curves confirmed by GC-MS. Cyclohexylamine, 6-hydroxyhexanoate and adipic acid were all derivatised with BSTFA + 1% TMCS to 10 g/L and diluted in Ethyl Acetate and again were confirmed by GC-MS. However 6-ACA would not dilute in this way and would produce two peaks at <1 g/L and three peaks at 1 g/L. The peaks followed a pattern of dilution e.g. reduced size at smaller concentrations and corresponded to the following; 6.94 minutes was identified as Trimethylsilyl 6-(acetylamino)hexanoate, 7.49 minutes was unidentified and only at 1 g/L was a small 6-ACA peak seen at 8.1 minutes. At no other concentration was this seen (Figure 6.6).

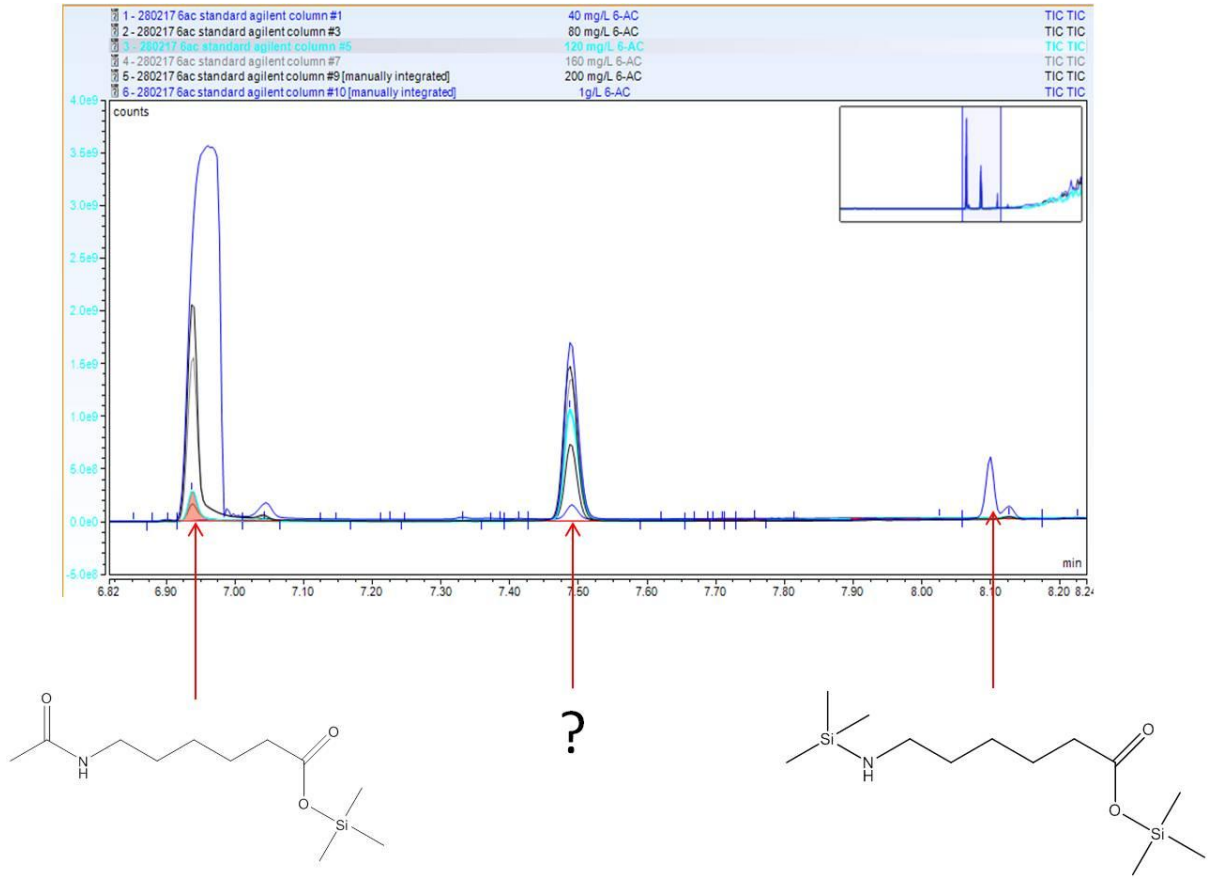



**Figure 6.6 Three peaks identified from 99% pure 6-ACA derivatised and diluted into lower concentrations.**

Given that the only solvent that could be found that 6-ACA is soluble in is water, a dilution series was made and speed vacuumed  until dry. This was then reconstituted into the original volume in ethyl acetate and BSTFA + 1% TMCS and derivatised. However, the

typical chromatogram was seen in Figure 6.7. Samples were also left for longer (14h) to derivatise as well as at 50:50 with Ethyl Acetate and BSTFA but this was also unsuccessful. A wide array of peaks were seen which couldn't be accounted for with signals strong enough to indicate that they are not trace quantities (Figure 6.7).

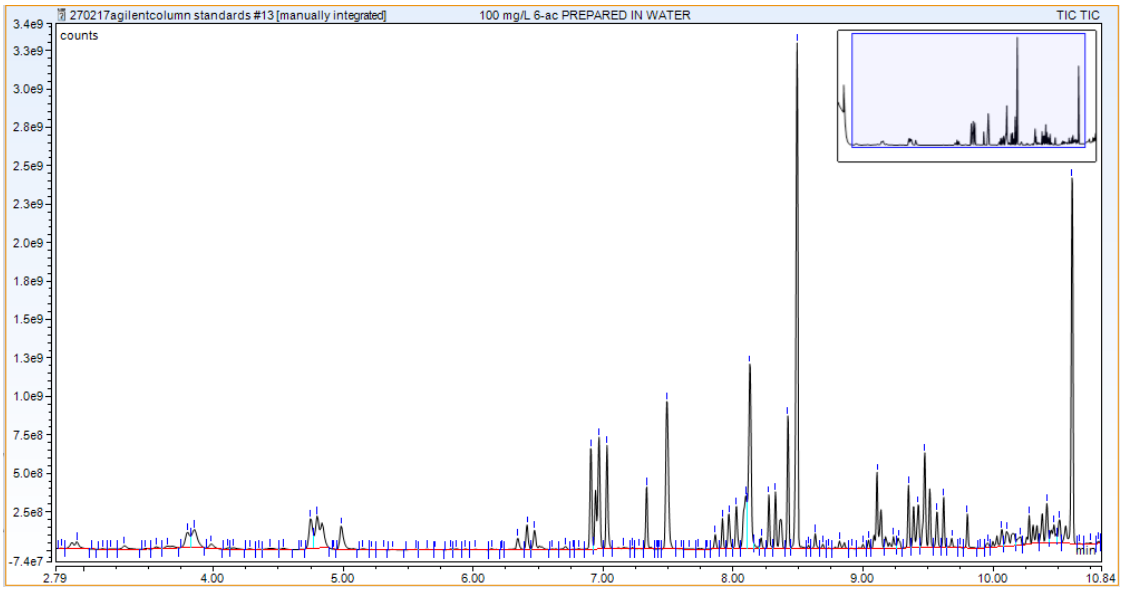



**Figure 6.7 6-ACA prepared in water and diluted out produced this typical pattern after drying and derivatising.**

Since 6-ACA was detected at the end of a fermentation in complex media but not with the standard it's possible that the acyl transfer from the Ethyl Acetate gets mopped up by more attractive amino acids or there is a potential pH influence. A set of three cultures of pQR1553 were grown, but not induced, for 18 hours in conical flasks in LB broth. 6-ACA standard was added to each to make a final concentration of 100 mg/L, 80 mg/L and 60 mg/L and left for 1 hour as a mimic for bioproduced 6-ACA. The cells were lysed, processed in the same way as in reaction cultures and this produced a linear concentration curve which was confirmed by GC-MS (see Figure 8.8 in the appendix for the graph).

### 6.2.3 Operon clones supplied with 10 mM cyclohexanol successfully produce 6-ACA

Now that a GC-MS method and a sample processing method have been developed the clones were more thoroughly investigated for 6-ACA bioproduction. Being water soluble, initial investigations were made to determine if 6-ACA was cell permeable and if so, what is the degree to which 6-ACA is partitioned between the media and intracellular environments. Previously cultures were grown supplemented with 200 mM cyclohexanol and subsequently showed a very small peak of 6-ACA which wasn't quantified however a large peak of cyclohexanol was noticed which indicates that there may be some substrate inhibition at that concentration. Therefore 10 mM was chosen as a significant reduction in substrate concentration and was spiked 2 hours post-induction. Only clones pQR1553 - 1556 were investigated at this stage because *MLUT*_00902 was proving a lot more problematic to clone. The cell cultures were then sampled and either sonicated followed by centrifugation or only centrifuged with no cell disruption. After particulate removal, the supernatant of both samples was then split into two fractions; one which was dried and then derivatised whilst the other fraction was mixed vigorously with Ethyl Acetate, centrifuged and then the solvent phase was analysed by GC-MS. Figure 6.8shows that in all but pQR1554, the cell lysate (CL) produced more 6-ACA than was in the supernatent (SN) alone. This is especially pronounced in pQR1553 which released 38 mg/L from the intracellular compartment. pQR1554 was approximately equal however the reason for this is not clear.

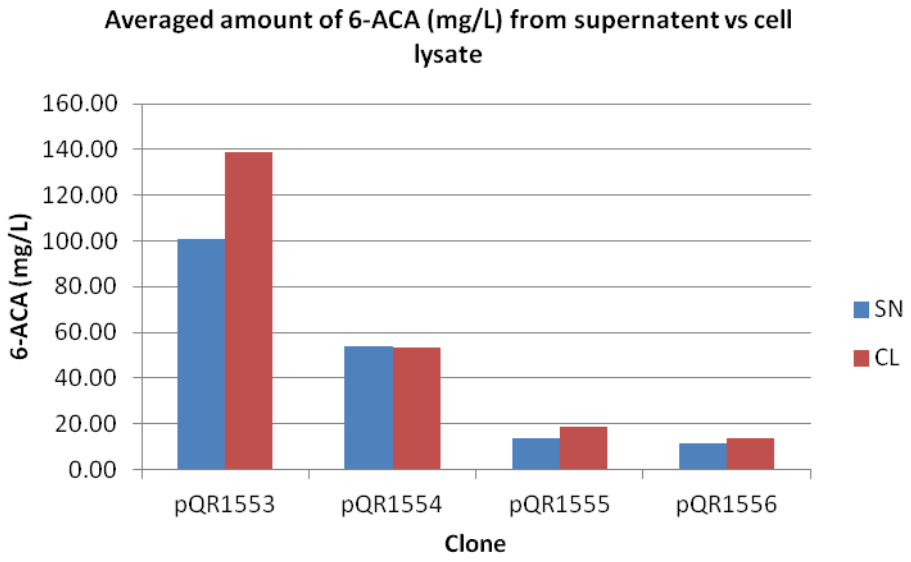



**Figure 6.8 Graph showing amounts of 6-ACA recovered from supernatent (SN) only compared to when the cells are lysed (CL) which is a meansure of the total of intracellular and extracellular 6-ACA. The operon clones used at this time did not include pQR1557 since this was in the process of being assembled. The clones were grown on 10 mM cyclohexanol for 16 hours.**

When looking at the whole intermediate profile of clones pQR1553 - 1557 (excluing 6-oxohexanoate; Figure 6.9) pQR1553 was identified as the largest producer of 6-ACA averaging 138.7 mg/L (operon constructed of natural RBSs and *PP5182*). Importantly it can be concluded that the RBSs have had significant impact on the performance of the pathway. pQR1554, having the same genes as pQR1553 but the RBSs are altered to *AGGAG* at position -7, achieved only 53.18 mg/L; a significant difference of 85.52 mg/L. pQR1557 only produced 86.29 mg/L and since in all other constructs the operons regulated by natural RBSs were shown to be the best produces, the artificial RBSs version containing *MLUT_00902* was not assembled. This was because of the difficulty of creating pQR1557 combined with a likelihood of no significant increase in performance was deemed to be of little benefit. No cyclohexanone was detected with pQR1557 however 303.87 mg/L of 6-hydroxyhexanoate was found. 6-hydroxyhexanoate was also present in all other clones in amounts greater than 127.57 mg/L. This validates the isolated enzyme activity of

6-hydroxyhexanoate dehydrogenase which highlighted its potentially low activity *in vivo*. Cyclohexanone was also only found in small amounts in pQR1553 compared to the other clones (averaging 12 mg/L). Interestingly all clones had large amounts (>54.86 mg/L) of adipic acid, the largest of which was in pQR1557 (157.16 mg/L); this was largely predictable based on the knowledge that TAm reactions are relatively slow compared to the other enzymes in the pathway as well as the abundance of aldehyde dehydrogenases in *E.coli*. Furthermore, the low or absent quantities of cyclohexylamine would seem to confirm the viability of the TAm screen results since the top three TAms' were chosen based on their low activity on 6-oxohexanoate. When looking at the molar amounts produced, only a total of 2.99 mM of the 10 mM fed could be accounted for in pQR1553; 1.06 mM of that was 6-ACA and 1.14 mM of 6-hydroxyhexanoate. In fact the highest combined molar amounts of intermediates was only 4.03 mM from pQR1557. Since the solvent extraction step for the first three molecules is immediate it is unlikely that there is great losses here. Instead, in the above experiments the lysed cell supernatant was dried in a drying oven at 75$^{o}$C overnight. This temperature and the extended time at which 6-ACA would spend dissolved in a slowly reducing volume could increase the likelihood of cyclisation of the amine and carboxylic terminals reacting together to form caprolactam. The method of drying samples was then investigated to identify the impact on 6-ACA production.

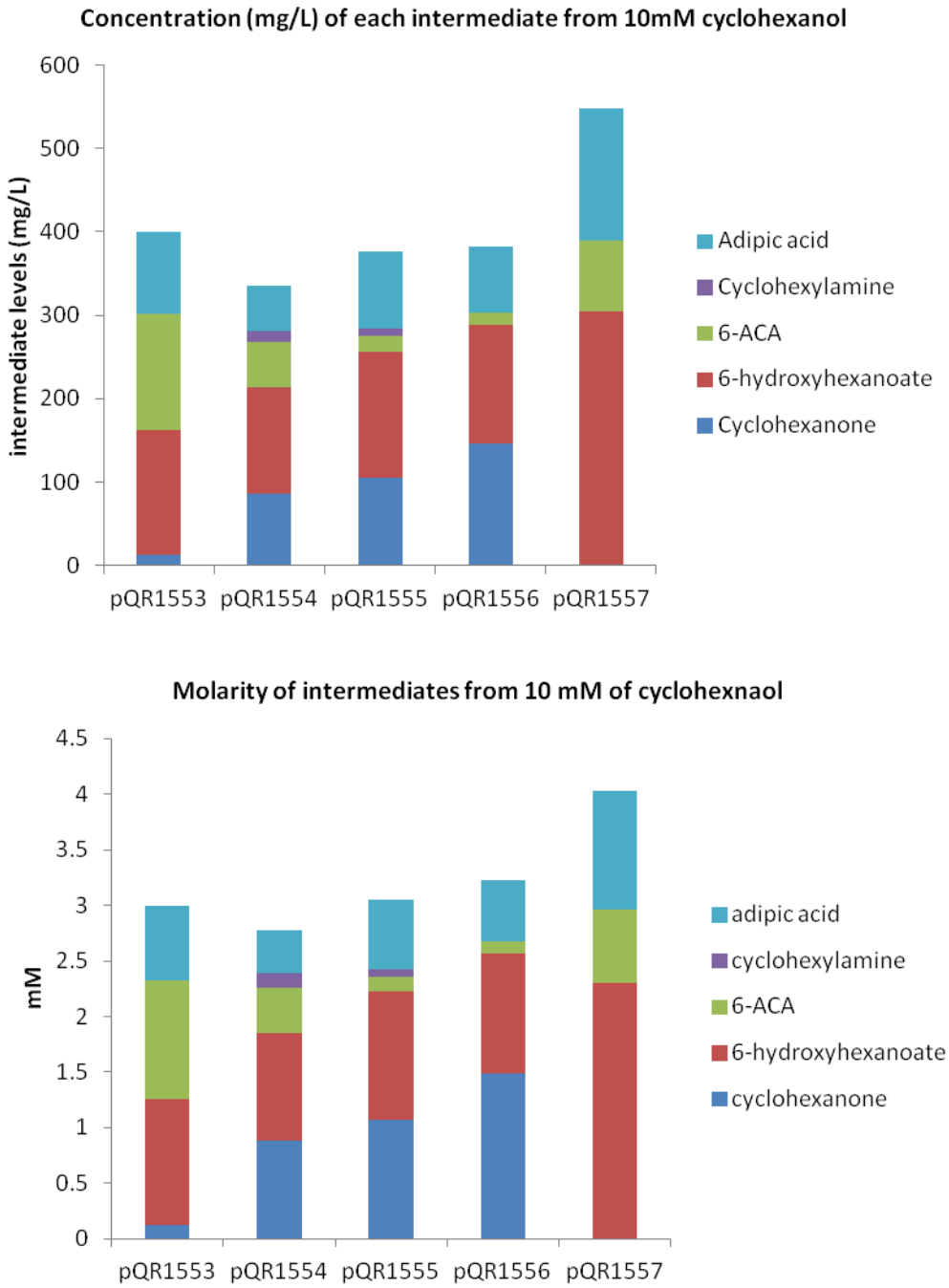
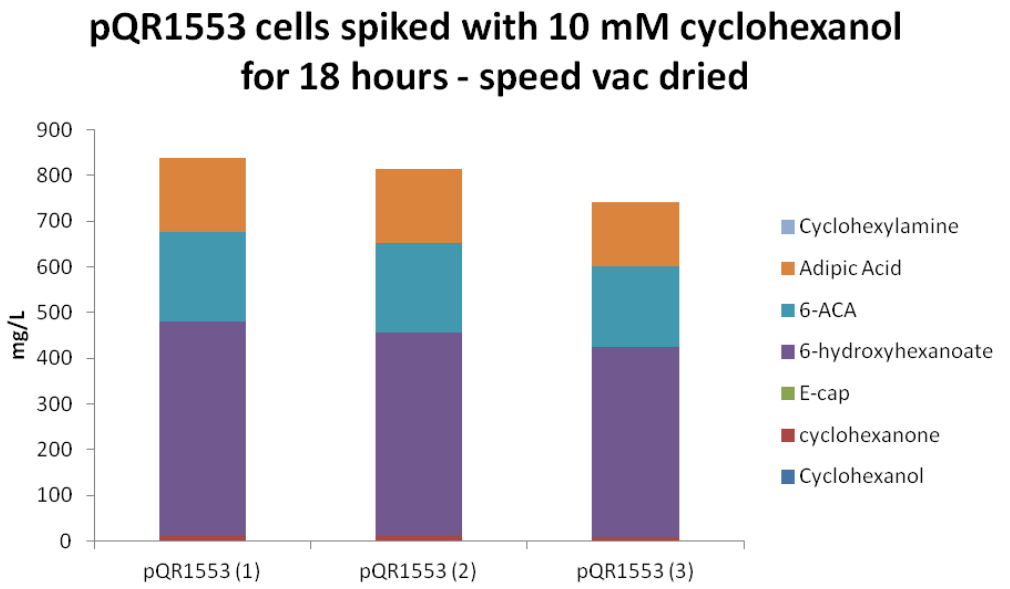



**Figure 6.9 Graphs showing intermediate concentrations in mg/L and molar amounts produced after 18 hours incubation with 10 mM cyclohexanol.**

### 6.2.4 The speed at which samples are dried post-sonication impacts upon detectability of 6-ACA

To test if the method of drying lysed samples had an impact on process yield only pQR1553 was used since it had been identified as the best producing clone. Previously a drying oven set to 75$^{o}$C had been used to dry samples overnight however as previously explained this could potentially be a source of problems with 6-ACA detectability. Triplicate cultures from separate colonies were grown and spiked with 10 mM cyclohexanol and left for 18 hours at 37$^{o}$C shaking. After this samples were taken in the same way as previously, the cells were disrupted, solid particulates removed and were subsequently dried in a speed vac until completely dry. This processes typically took 2-3 hours and was done at room temperature. The solid pellet was then mashed and broken up into fine powder using a pestle and mortar fashion in ethyl acetate and BSTFA. Figure 6.10 shows the concentrations of intermediates detected by GC-MS. Overall the total molar amounts of intermediates increased to an average of 5.97 mM, nearly double that which was previously detected. The amount of 6-ACA now observed averaged at 189.54 mg/L. Cyclohexanone obviously remained the same however a significant finding was the large increase from 150.28 mg/L to 443.51 mg/L of 6-hydroxyhexanoate. This could also be due to the previous cyclisation of 6-hydroxyhexanoate into caprolactone and was not measured at the time. Adipic acid also increased from an average of 98.85 mg/L to 154.8 mg/L although the reason for this is unclear. Possible reasons for this include that a newer column was used in this second round of 10 mM cyclohexanol growths which might account for some increase in sensitivity however these results suggest that the speed and possibly temperature at which the samples are dried has a drastic impact on measured intermediate concentrations.



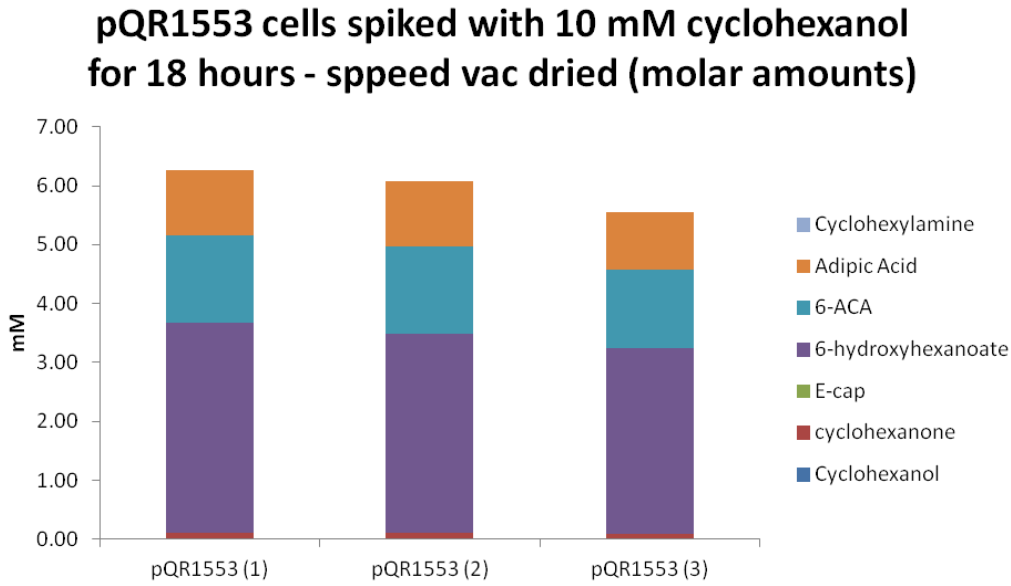
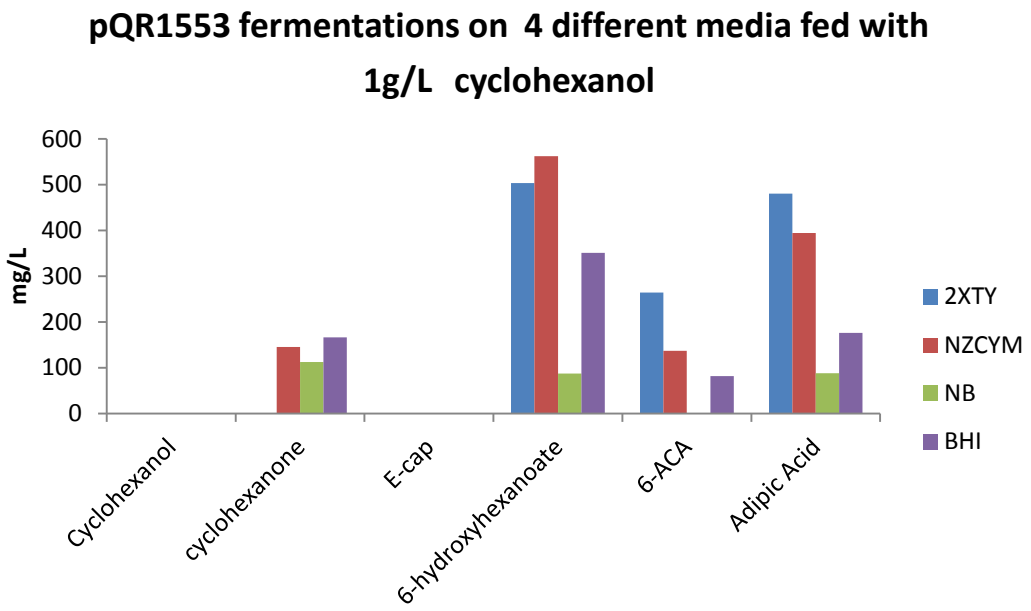




**Figure 6.10 Intermediate concentrations for pQR1553 cells grown with 10 mM cyclohexanol and dried via speed vac for 2-3 hours until complete; (mg/L above, mM concentration below)**

### 6.2.5 Media screen on 10 mM cyclohexanol with pQR1553 cells increases yields significantly

Cell culture media is known to affect protein expression, alter metabolism and nutrient profiles and ultimately affect cell density and fermentation performance (Sivashanmugam *et al.*, 2009). It may also alter cell membrane protein and the permeability of a cell. These factors need to be optimised for efficient production of any bioproduct however generating a defined media is a very arduous process. Instead a selection of four different complex media were used to grown pQR1553 cells and spiked with 10 mM cyclohexanol for comparison with LB media previously used. Intermediate compound concentrations were found to vary significantly between media. 2xTY media produced 264.48 mg/L of 6-ACA compared to none from NB. Media choice also impacted on the amount of cyclohexanone present at the end of the growths and in 2xTY none was found however the next greatest producing media of 6-ACA, NZCYM, was found to have 145.42 mg/L of cyclohexanone (Figure 6.11). Possible reasons for this include the reduced expression of *ChnB* on NZCYM or an altered metabolism affecting the availability of NADPH. It is clear that in the two best performing media (2xTY and NZCYM), significant bottlenecks in the pathway are still the abundance of 6-hydroxyhexanoate and adipic acid. These different media have different amino acid concentrations to LB and even though some may achieve a higher OD, this doesn't necessarily correspond to a high yielding process. Interestingly, the total molar concentration of the intermediates in 2xTY and NZCYM was 9.12 mM and 9.48 mM respectively. Compared to an average of 5.97 mM on LB, this represents a significant increase in molar conversion from 10 mM cyclohexanol.



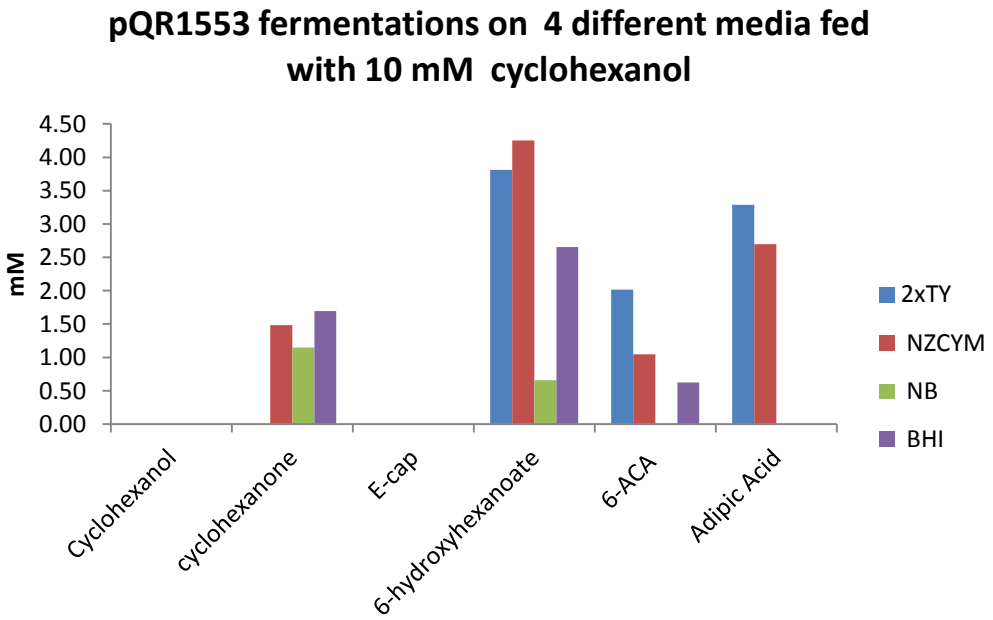
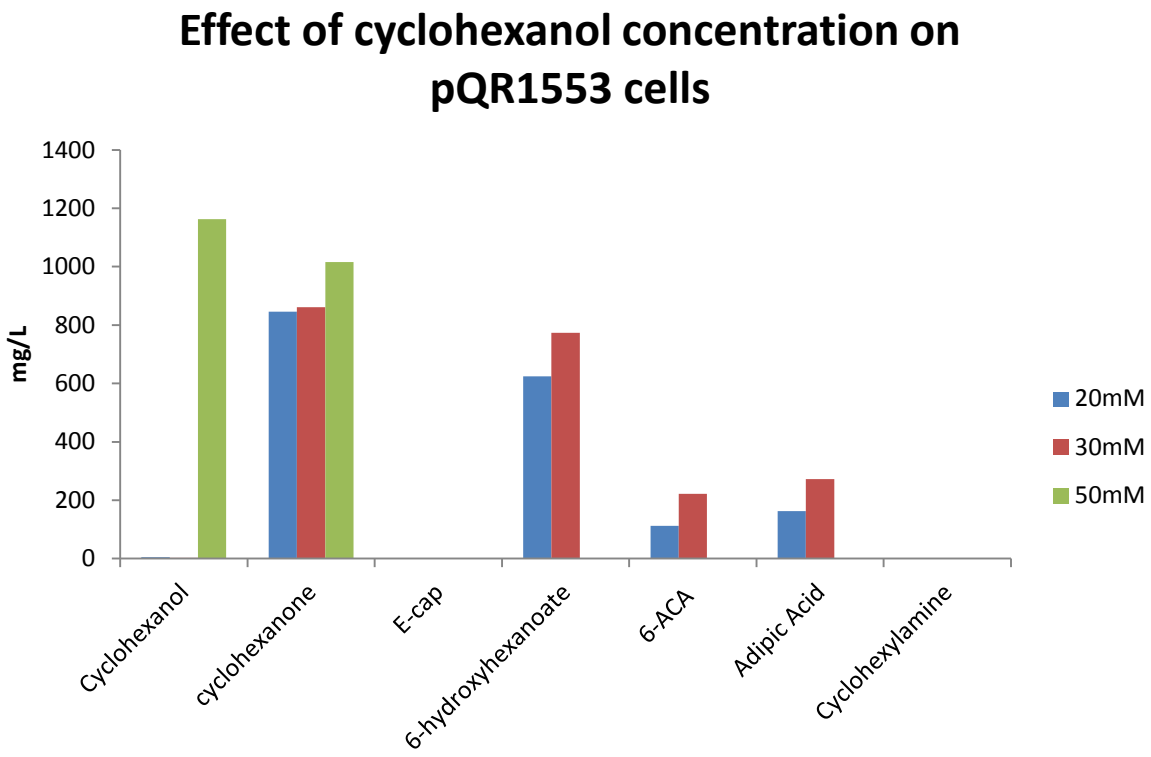




**Figure 6.11 The intermediate profile in mg/L above and molar concentrations from pQR1553 grown on different media and spiked with 10 mM cyclohexanol 2 hours after induction. See section 2.7 in the Materials and Methods for details on the culture conditions.**

### 6.2.6 Cyclohexanol concentration screen on pQR1553 in LB media

In order to get an idea of the range at which cyclohexanol can be spiked at, three increasing concentrations were tested on pQR1553 cells grown in LB media. Cultures were grown in the same way as before and fed 20 mM, 30 mM and 50 mM bolus of cyclohexanol. Figure 6.12 shows that 30 mM of cyclohexanol produces 222.3 mg/L (1.69 mM)of 6-ACA compared to 10 mM producing 189.54 mg/L (1.45 mM). For reasons not yet known, 20 mM only produced 112.5 mg/L (0.86 mM) but produced slightly more adipic acid (162.27 mg/L (1.11 mM) on 20 mM cyclohexanol and 154.8 mg/L (1.06 mM) on 10 mM cyclohexanol). At all concentrations however the amount of cyclohexanol increased by a minimum of 81 .73 fold by doubling the amount of cyclohexanol from 10 mM to 20 mM. No cyclohexylamine was found in any sample. 6-hydroxyhexanoate was found at 624.54 mg/L (4.73 mM) at 20 mM cyclohexanol and 773.07 mg/L (5.85 mM) on 30 mM cyclohexanol, a sharp increase from the 443.51 mg/L (3.36 mM) when grown on 10 mM cyclohexanol. Interestingly, when spiked with 50 mM only very large amounts of cyclohexanol (1.162 g/L, 11.6 mM) and cyclohexanone (1.015 g/L, 10.3 mM) were found. When total molar amounts of intermediates are calculated it was found that proportionate to the starting concentration of cyclohexanol, fewer moles were accounted for at the end of fermentation as the concentration of cyclohexanol increased, i.e. 15.36 mM of the 20 mM cyclohexanol was accounted for whereas only 21.96 mM of the 50 mM cyclohexanol substrate was accounted for. Although it was found that 30 mM cyclohexanol increased 6-ACA production on LB, there was proportionally more cyclohexanone and 6-hydroxyhexanoate generated by this increase in cyclohexanol; 8.67 mM increase in cyclohexanone and 2.49 mM increase in 6-hydroxyhexanoate compared to only a 0.26 mM increase in 6-ACA production.



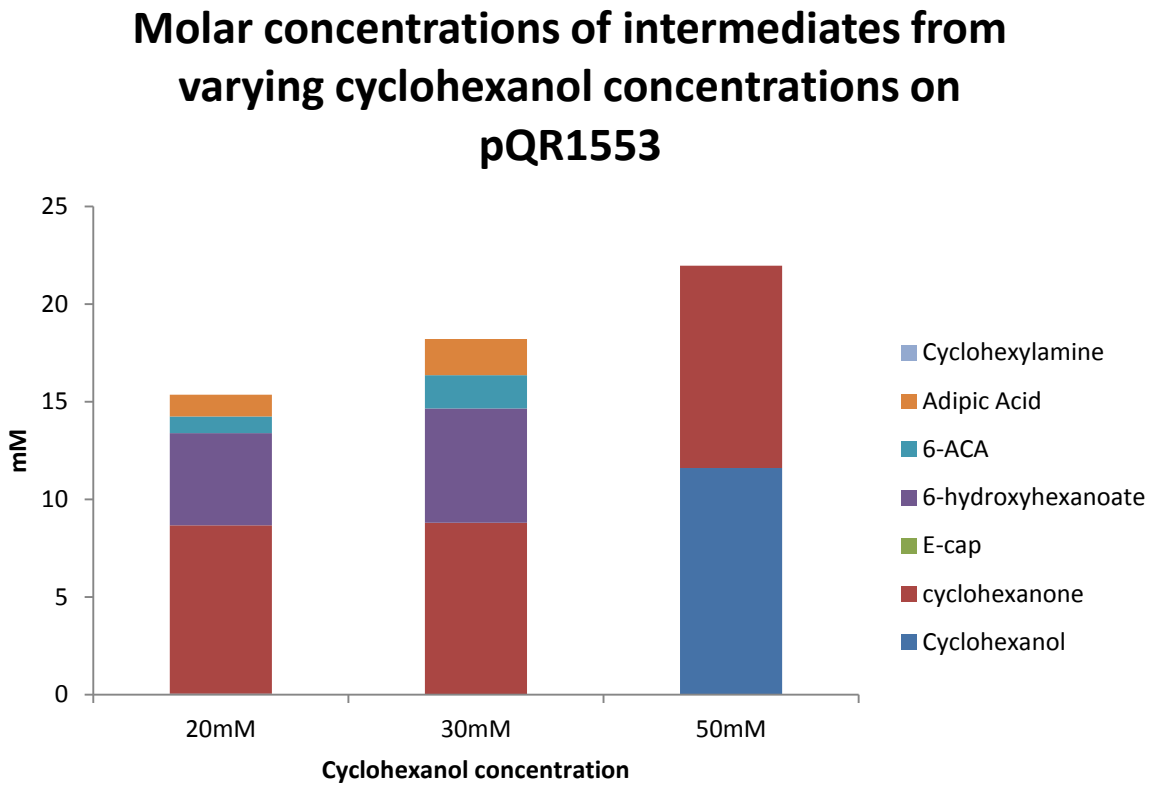




**Figure 6.12 Graph of pathway intermediates measured when pQR1553 is grown on LB media spiked with different concentrations of cyclohexanol. Below is a graph showing total molar concentrations of the pathway intermediates when grown with different cyclohexanol concentrations. (see Materials and Methods section 2.7)**

### 6.2.7 Growth and pH are affected by the concentration of cyclohexanol

During previous experiments it was observed that higher cyclohexanol concentrations produced smaller cell pellets during processing. Although 30 mM cyclohexanol has previously been shown to be the most productive, Figure 6.13 shows that when spiked with 30 mM and 50 mM concentrations of cyclohexanol growth is impaired, albeit to the same degree, whereas OD was higher when grown on 10 mM but conversely 6-ACA yield was lower. The effect 6-ACA has on the pH of a culture is unclear although an impact of the pathway can be seen. The cells were induced at 0.6-0.7 OD, and spiked with cyclohexanol at 2 hours post induction. It can be seen that from this point on, the previously matching pH curves deviate from each other; 10 mM produced a final pH of 8.39, 30 mM at 7.07 and 50 mM at 7.83. If no or very little 6-ACA is produced by 50 mM, the pH of the media will be buffered by other salts or amino acids in the media. At 30 mM, producing more 6-ACA lowered the pH to almost neutral whilst slightly less is produced at 10 mM a more basic pH is observed. 6-ACA is a lysine analogue which is a basic amino acid, that is it binds $H^+$ ions onto the ε-$NH_2$ group, becomes protonated and therefore pushes the solution towards an alkaline pH. It could be expected that the more 6-ACA is produced, the higher the pH of the culture would be and so it would be reasonable to assume that cultures grown on 30 mM cyclohexanol would be more basic. The reason why it is in fact more acidic than even cultures which don't produce any 6-ACA can be explained by the effect of the copious amounts of 6-hydroxyhexanoic acid which are produced at 30 mM compared to 10 mM cyclohexanol i.e. there's only a 0.26 mM increase in 6-ACA compared to a 2.49 mM increase in 6-hydroxyhexanoate. This large increase in 6-hydroxyhexanoate could result in the lowering of the pH and offset any increase in the pH by the relatively small increase in 6-ACA production. Therefore higher cyclohexanol feeding concentrations reduce the culture OD but there is a balancing act between pathway intermediates as to its effect on pH.

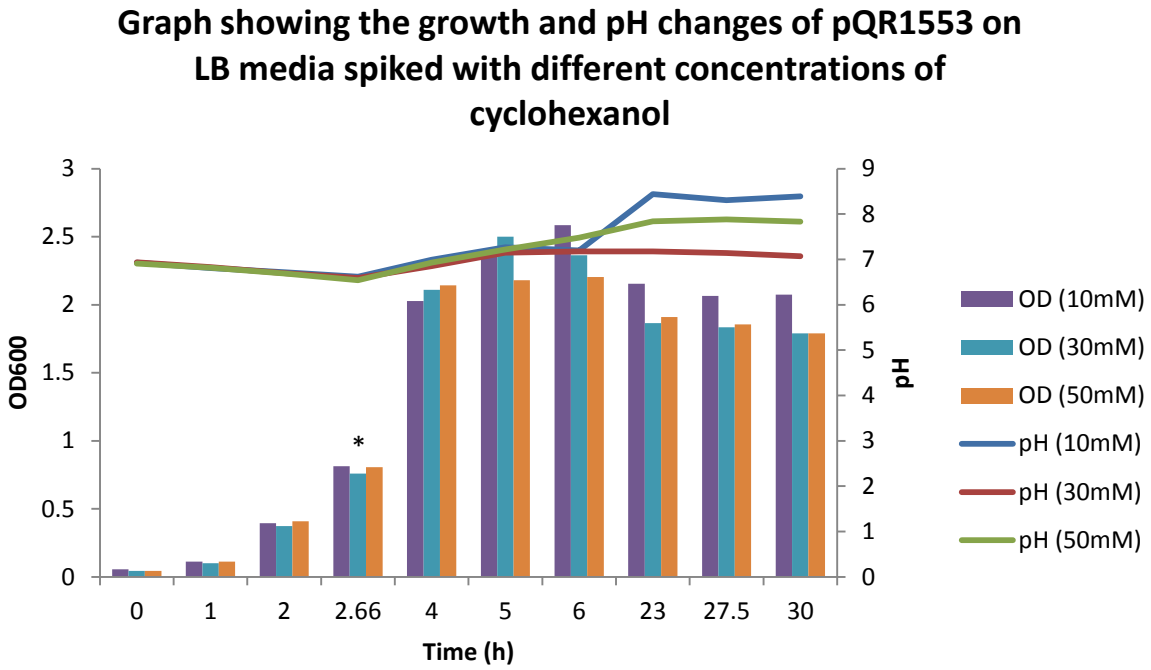



**Figure 6.13 pH and growth curves of pQR1554 on increasing amounts of cyclohexanol. Cultures were inocculated to an OD of 0.1 and induced at 0.6-0.7 OD. pH measurements were taken using a pH probe at the same time as an OD measurement was taken. See Materials and Methods sections 2.5.4**

## 6.3 Cloning a second *ChnD* under the control of a second T7 promoter and terminator into pQR1553

Results shown in chapter 3.4.3, 5.2.3 and onwards have generated evidence to support the claim that 6-hydroxyhexanoate dehydrogenase is an enzyme of low activity. This has led to a significant bottleneck in the pathway representing a maximum of 56.3% of total detected molar intermediates, when spiked with 10 mM cyclohexanol in LB media and a minimum of 30.7% when spike with 20 mM cyclohexanol in LB media. In order to try and reduce this, a significant increase in expression of the rate limiting enzyme, *ChnD* was required. To do this, one approach would be to clone a second copy into the plasmid of pQR1553, however if it was included in the operon directly after *PP5182* (i.e. the 6th gene in the operon) the

impact is likely to be much lower than if it was under control from its own promoter and terminator. Therefore a plasmid clone carrying the operon from pQR1553 as well as a second copy of the T7 promoter and terminator at the opposite side of the plasmid was required.

### 6.3.1 pQR1559 assembly

After looking at the pQR1553 sequence a suitable location to add another promoter, operator and terminator was identified. A region 343 bp upstream of the ROP gene was chosen and primers were designed to linearise pQR1553 by inverse PCR from this point and the for primer from 34 bp down from this which removes an area of large GC content. *AvrII* and *SalI* restriction sites were added to the ends of the linearised pQR1553 for restriction cloning once the *ChnD* fragment was created (Figure 6.14). Next, using the nested *ChnD* fragment generated from 1.1, primers were designed to add on an operator, a T7 promoter and and *AvrII* site from the natural RBS of *ChnD*. Since these features are small (50bp in total plus 6 random bp before the *AvrII* site which will get cleaved before ligation), they were successfully amplified (Figure 6.15). No terminator was added to the fragment at this point only a simple reverse primer from *ChnD* was used. Both of these fragments were gel purified and stored at -20$^{o}$C.

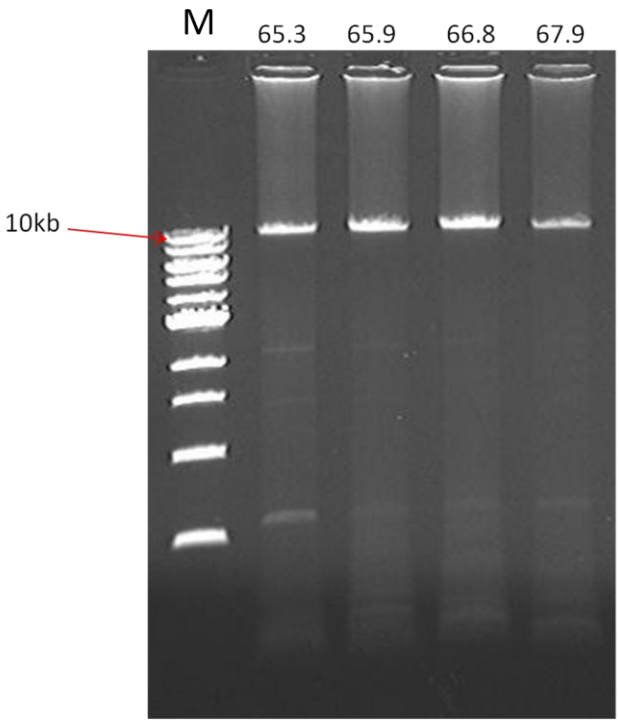



**Figure 6.14 Inverse PCR of pQR1553 adding *AvrII* and *SalI* restriction sites. The Annealing temperature is shown above each sample gel lane. Slightly more off target effects are seen at the lowest temperature compared to the highest. Marker lane (M) is a 1 kb ladder (NEB, USA).**

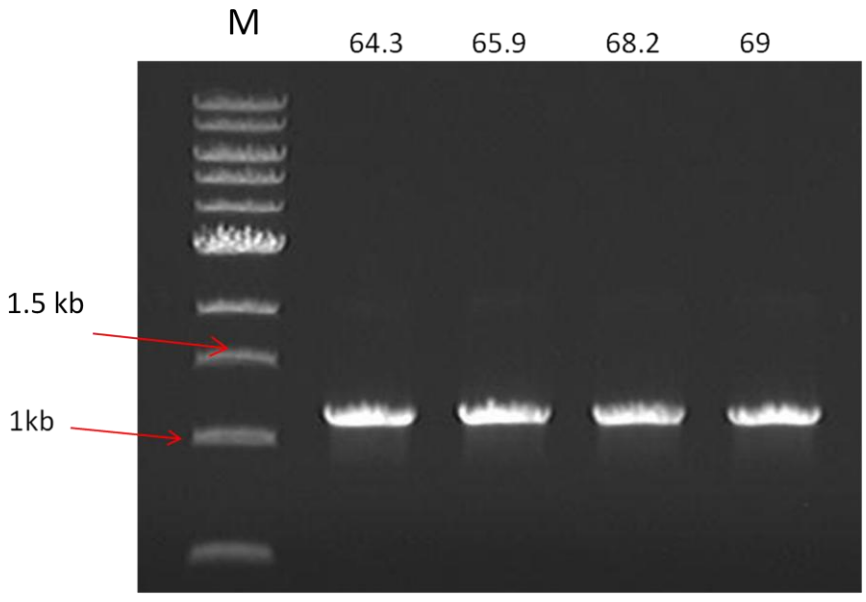



**Figure 6.15 PCR on the nested *ChnD* including the natural RBS adding on a T7 promoter and operator and *AvrII* site upstream of the gene. Annealing temperatures are shown above the lanes. Marker (M) lane is a 1 kb ladder (NEB,USA).**

This *ChnD* fragment was then used as a template to add a terminator at the 3' end of the gene. The T7 terminator is 48 bp and so a primer was designed to incorporate this in the next round of PCR. Figure 6.16 shows that this was successful.

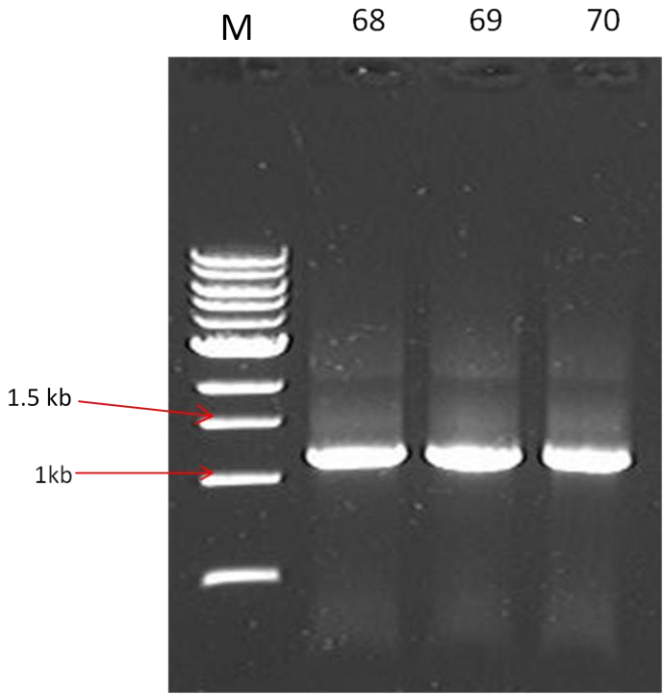



**Figure 6.16 *ChnD* incorporating *AvrII*, T7 promoter and operator followed by the T7 terminator. Annealing temperatures are shown above each lane. Marker (M) is a 1kb ladder (NEB, USA).**

Next the PCR was gel extracted and moved into the final PCR stage in order to add a *SalI* restriction site to the end of the terminator sequence. Again this PCR was successful (Figure 6.17) and was gel extracted.

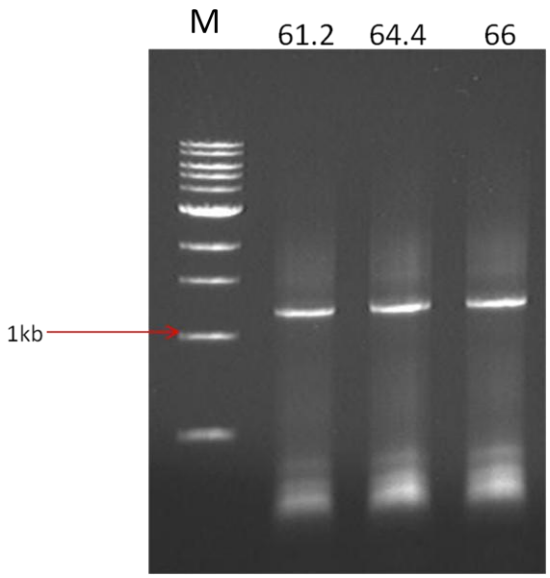



**Figure 6.17 The final PCR generating the complete *Avr*II, T7 promoter, operator, *ChnD*, terminator and *Sal*I fragment. Annealing temperatures are shown above each lane. Marker (M) is a 1kb ladder (NEB, USA)**

446.5 ng of PCR linearised pQR1553 and 780 ng of the new *ChnD* insert were digested by *Avr*II and *Sal*I for 4 hours followed by gel purification. The two fragments were ligated together at an insert molar excess 20 times that of the vector backbone and transformed into *E.coli* 10β cells (NEB, Massachusetts, USA). Five colonies were grown in 5 ml overnight cultures and miniprepped. Digestion of each colony with *Avr*I and *Sal*I for 4 hours confirmed the presence of two predicted bands; T7-operator-RBS-*ChnD*-terminator at 1.172 kb and an 10.938 kb band corresponding to the pQR1553 backbone (Figure 6.18). This clone was renamed pQR1559. Colony 4 was transformed into BL21 (DE3) and samples were grown for SDS-PAGE. A comparison between pQR1553 and pQR1559 shows there maybe some slight increase in expression of all proteins which may in fact indicate that simply more total protein was loaded since it was only loaded by volume and not equal amounts of total protein. Nevertheless 5 bands at expected molecular weights were confirmed (Figure 6.19).

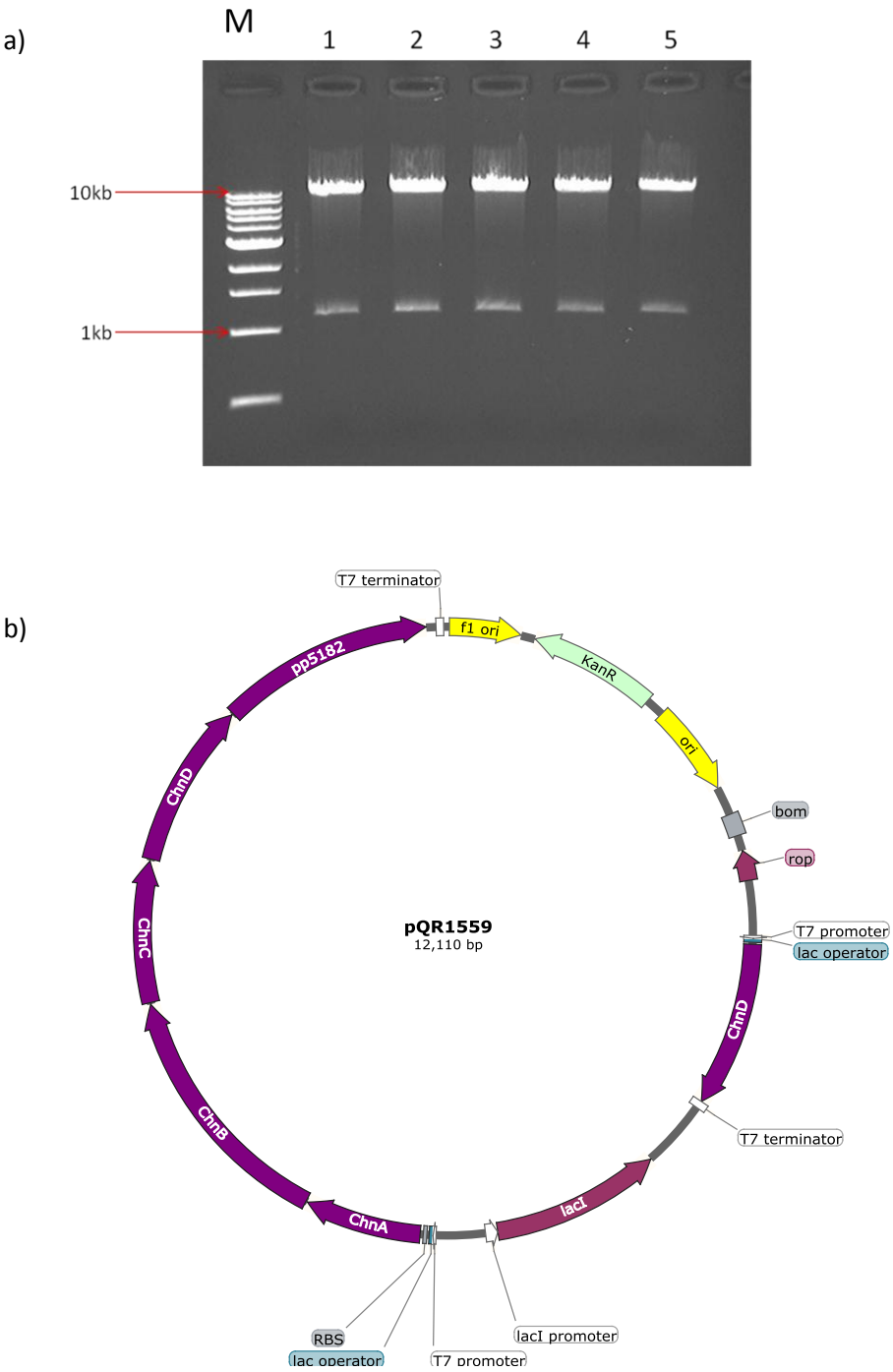



**Figure 6.18 (a) Is the *AvrII* and *SalI* digestion of 5 plasmid minipreps (lanes 1-5) from the ligation reaction of a second *ChnD* with a seperate promoter, operator and terminator. (b) Plasmid map of pQR1559. Marker (M) is a 1 kb ladder (NEB, USA).**

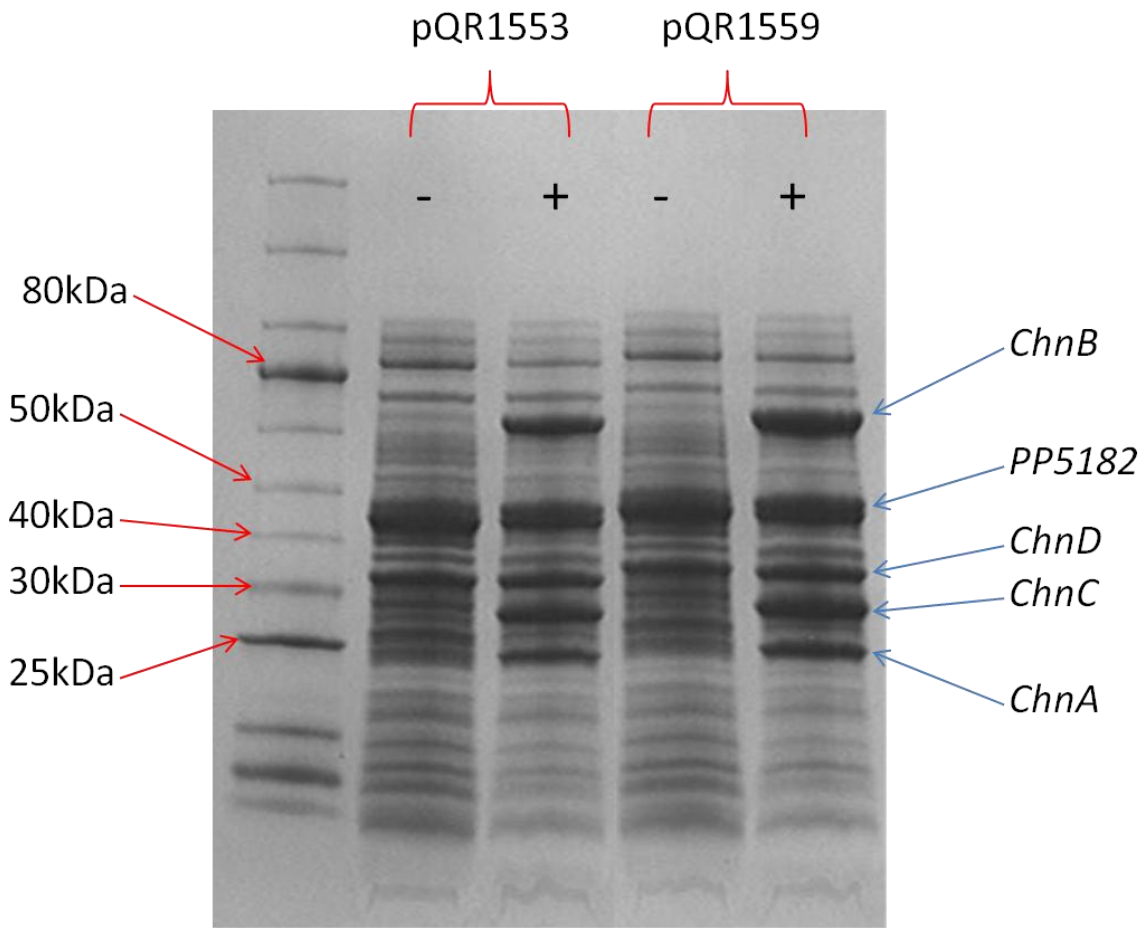



**Figure 6.19 pQR1553 and pQR1559 on an SDS PAGE. Samples taken from the same volume of culture and the same volume loaded onto the gel. Five clear bands can be seen at the expected molecular weights.**

### 6.3.2 pQR1559 performance on 10 mM cyclohexanol compared to pQR1553

When pQR1553 was grown on LB and spiked with a bolus of cyclohexanol to a working concentration of 10 mM, samples were taken 18 hours later. When analysed by GC-MS, the intermediate profile was different to that of pQR1553 when grown under identical conditions . This indicates the functional expression of *ChnD* and possibly an alteration in the levels of expression of the other enzymes. Expected results were that 6-hydroxyhexanoate would decrease and 6-ACA concentration would increase. Whilst 6-ACA levels increased by only 0.2 mM, conversely all other intermediates also increased, particularly 6-hydroxyhexanoate by 1.21 mM (159 mg/L) (Figure 6.20). A possible reason for this is that the demand from another promoter and a 6$^{th}$ gene placed a strain on the

host cells and either the amino acids or available cofactor profile was altered. This would cause the enzymes to have lower activity and consequently alter the pathway intermediates. However another possibility could be that the flux through the pathway actually increased but the TAm, being relatively slow, couldn't process the 6-oxohexanoate at a rate which was fast enough to either avoid aldehyde dehydrogenation into adipic acid or a reverse reaction back into 6-hydroxyhexanoate. Adipic acid concentrations increased by 0.79 mM (115 mg/L) which could substantiate the second possibility since the protein expression profile didn't look to be drastically different. Also interestingly, only a small increase in cyclohexanone and no cyclohexanol or ε-caprolactone was found which suggests that the flux through the upstream enzyme steps weren't changed drastically. This then implies that the downstream gene expression profile wasn't changed either since they're in the same operon. Therefore the bottleneck could actually be the TAm.

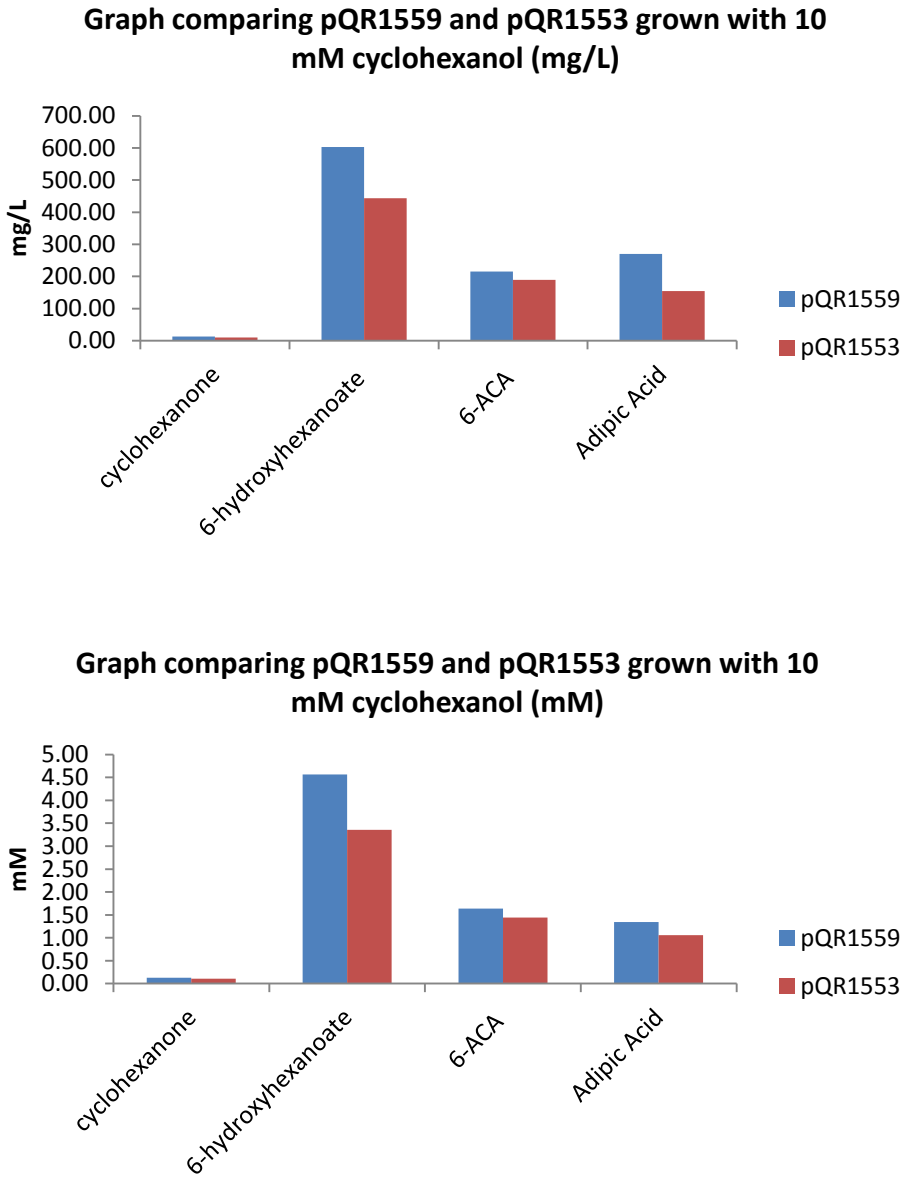



**Figure 6.20 The concentration in mg/L of intermediates (top graph) and the molar concentration (bottom graph) of the intermediates with a comparison to pQR1553. All metabolites increased in pQR1559 indicating a potential bottleneck with the TAm.**

### 6.3.3 pQR1559 performance on higher cyclohexanol concentrations and a small media screen

In contrast to pQR1553, when cyclohexanol is increased to both 20 mM and 30 mM, although the latter produces more of every intermediate than the former, both are less productive than when grown on 10 mM cyclohexanol. When grown on 30 mM, only 151.24 mg/L of 6-ACA was achieved compared to 215 mg/L 6-ACA when grown on 10 mM cyclohexanol. No solvent extraction phase was ran with these experiments because the samples were accidentally disposed of by a colleague (Figure 6.21). Because of the disappointing result with this clone it was not repeated.

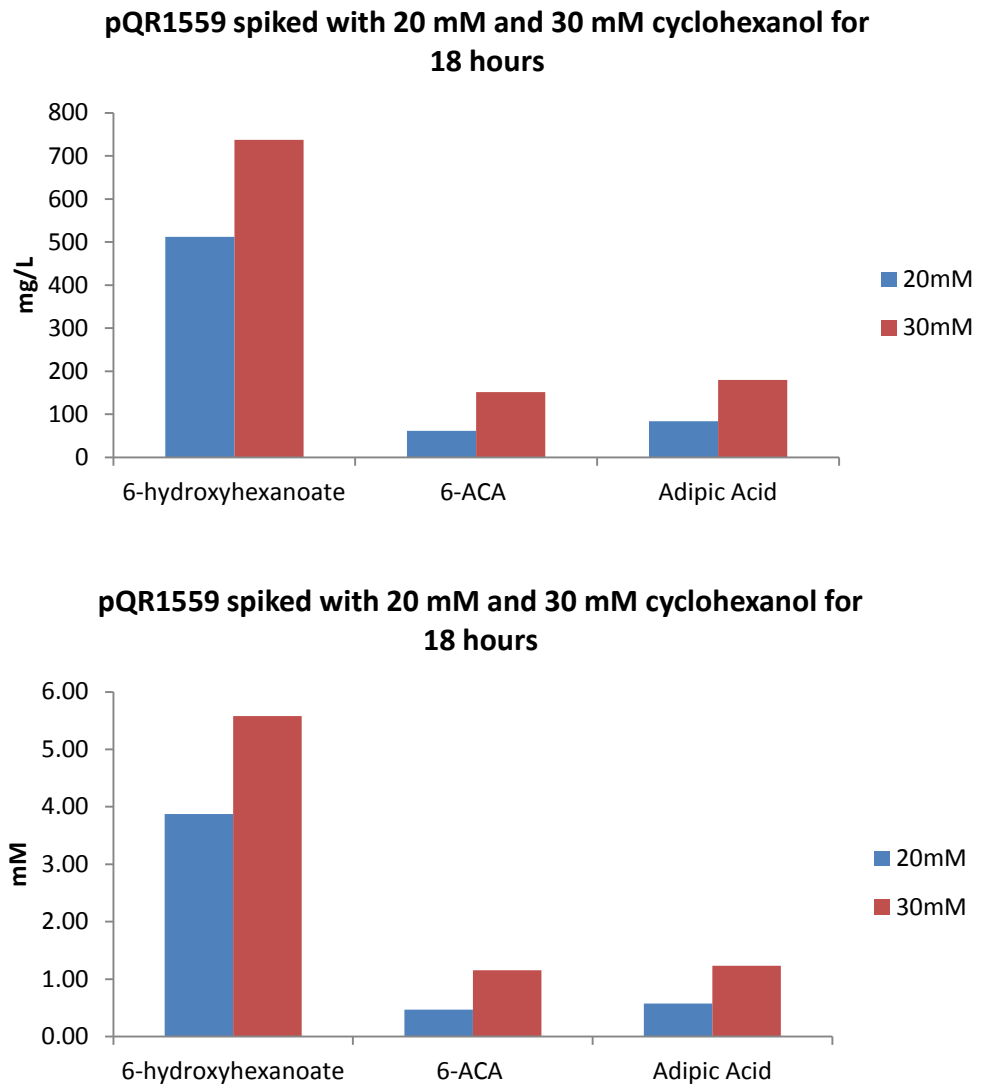
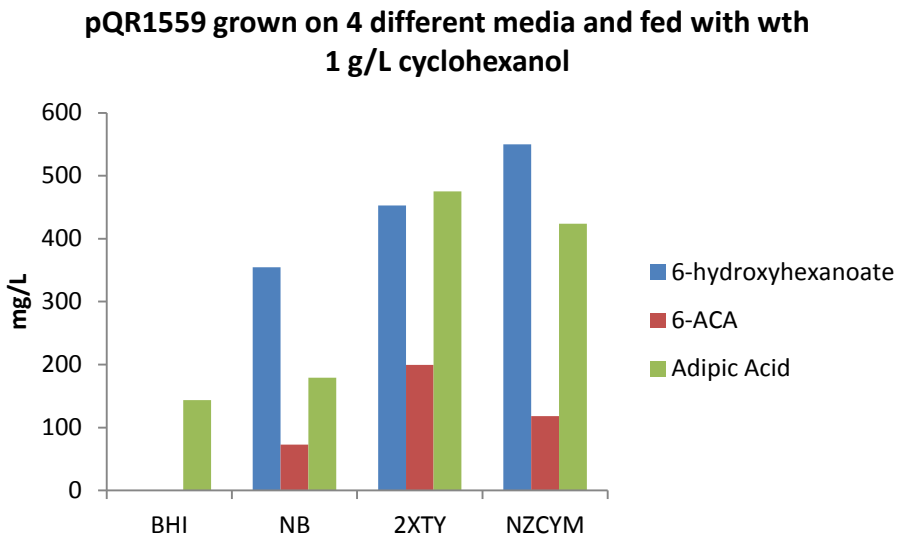



**Figure 6.21 The graphs show the downstream intermediate concentration achieved when grown on 20 mM and 30 mM cyclohexanol in mg/L (above) and molar concentrations (below).**

Furthermore pQR1559 performed less well when spiked with 10 mM cyclohexanol on the four different media compared to pQR1553. The trend was the same in terms of the order which produced the most 6-ACA with 2xTY performing best and BHI the least however pQR1559 only produced 199.5 mg/L on 2xTY whereas pQR1553 produced 264.48 mg/L (2.016mM) (Figure 6.22).

a)




b)

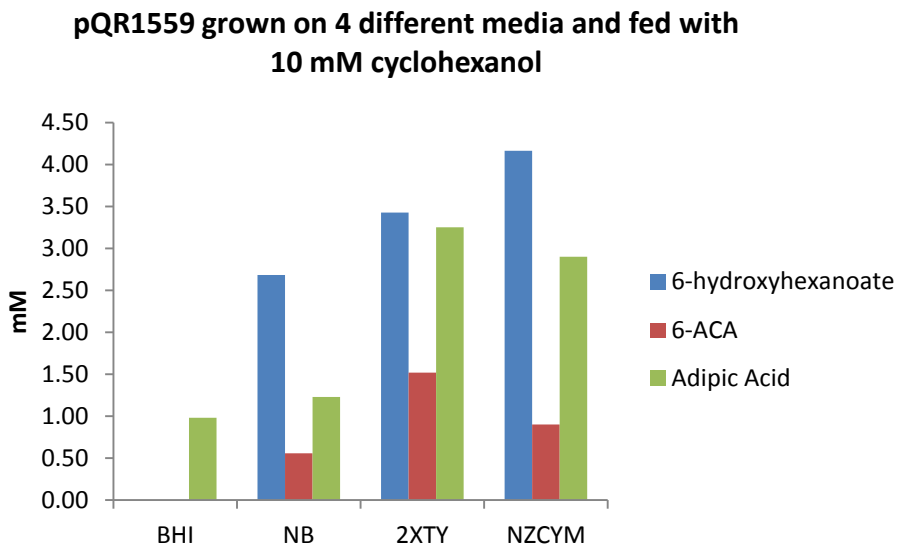



**Figure 6.22 pQR1559 was grown in 4 different media and fed with 10 mM cyclohexanol. The cultures were sampled 18 hours later. The intermediate concentrations achieved in mg/L (a) and below is the intermediate molar concentrations (b).**

# Chapter 7: Discussion

# 7 Discussion

## 7.1 The *Acinetobacter genes:* the first stages of operon assembly and individual characteristics

Genomic DNA preparations of *Acinetobacter calcoaceticus* NCIMB 9871 were used to amplify nested fragments of *ChnA, ChnB, ChnC,* and *ChnD*. These were then used as templates for further rounds of amplifications. The ultimate goal was to have all genes required on a single plasmid so as to only rely on a single antibiotic resistance marker as well as to avoid difficulties in co-transformations and potentially causing excess stress on the cells from two plasmids. Assembling multiple genes in a single plasmid by standard restriction cloning would be very challenging, possibly requiring many iterations, where each gene would be double digested and all would be assembled with unique restrictions sites. However in the late 2000's, the advent of several cloning methods which utilise homologous overlaps used a combination of multiple enzymes or single enzymes working in harmony such as exonucleases, polymerases and ligases which allowed multiple genetic elements to be assembled quickly, easily and include directionality (Gibson *et al.*, 2009). Initially each gene was amplified to encode homologous overlaps, 15-25 bp long, as recommended for Gibson assembly®. Initial data showed a difficulty in assembling the 4 genes together and produced bands of varying sizes with no clear target band assembled. Further investigation showed that the fragments of *ChnC* and *ChnD* assembled into a band of expected size however *ChnA* and *ChnB* failed to do so. Excess assembly mix and time showed better results however overall it seemed unlikely to be able to assemble 5 fragments. Although rarely used in the literature, CPEC afforded the capability to use the same fragments designed for Gibson assembly® only in a cyclical, temperature profile

assembly method which is an extension of PCR but yields no amplification. This method proved to be very successful and very time efficient producing an operon in as little as a week from PCR to Cloning and SDS-PAGE verification. When the four *Acinetobacter* genes were assembled, a prominent expected band of 4.4 kb was seen. When made circular by adding the plasmid backbone the assembled DNA wouldn't migrate from the gel well. The reason for this is unknown however it's possible that at this stage the majority of DNA is entangled which prevent movement into the gel. When the well was excised and transformed, colonies were produced, the majority of which digested as expected, producing bands at 4.15 kb and 5.48 kb. Incorrect assembly is likely, to a certain degree, with CPEC and can be thought of in a similar way as mis-priming and off-target effects can be seen in a PCR. One clone was sequenced at the joining ends confirming the assembly of *ChnA* and *ChnD* and the SDS-PAGE showed that all four proteins were expressing at predicted molecular weights. Interestingly however after 23 hours of growth, CHMO ceases to express. Given that this is the second gene in the operon and the others aren't affected it is likely that the protein is expressing however is unstable and therefore degraded easily. Two operon versions were created, one harbouring natural RBS sites identified before each gene and one with an artificial RBS, both cloned easily and quickly using CPEC. The rationale behind this is that perhaps *Acinetobacter* has evolved and has finely tuned each gene expression to optimise flux through the pathway and so by using this, we can hypothetically achieve a maximum amount of 6-oxohexanoate funnelled through to the TAm. This is however assuming that *E.coli* BL21 (DE3) uses very similar RBS strength recognition.

Moving on from these preliminary operon versions, each gene was cloned individually into a pET29a vector and His-tagged to allow Ni-affinity purification. The main aim was to be able to add to the library of cloned enzymes in the group but also to allow each enzyme to be assayed individually for activity indications. The alternative to this was to feed in

cyclohexanol and identify the intermediates however the extraction protocols, sample preparation and analytical methods need considerable thought and at this early stage, the quickest way was to clone them individually. Since most use a readily available co-factor, a simple spectrophotometry screen would suffice. The only enzyme with no detectable co-factor requirements is the lactonase however activity was assayed by coupling with 6-hydroxyhexanoate dehydrogenase, supplying caprolactone as a substrate as well as $NAD^+$ showing that if the lactonase is active and producing 6-hydroxyhexanoate, the 6-hydroxyhexanoate dehydrogenase will be able to use that as a substrate and produce an absorbance change. A range of substrates were tested and results suggested that both CHMO and 6-hydroxyhexanoate dehydrogenase could potentially be bottlenecks in the pathway. CHMO exhibited substrate inhibition which is well documented and when tested as a standalone enzyme, 6-hydroxyhexanoate dehydrogenase showed little activity at all substrate ranges tested. However, the lactonase experiment showed it was clearly active however. These results were only indicative and were not meant to be absolute since their characteristics *in vivo* as well as the amount of protein expressed will be different. Since assembling the operon was initially thought to be likely to be the most problematic and time consuming part, that was completed first. That way if the enzymes showed no activity, alternatives could be found and substituted into the operon. Overall however, cyclohexanol dehydrogenase and the lactonase exhibited excellent activity profiles and were able to tolerate high concentrations of substrate. Even though the CHMO and 6-hydroxyhexanoate dehydrogenase were likely the weak links, no enzyme substitutions were made since they both showed activity and when working in the pathway, may behave differently when the dynamics of their respective products become substrates for the next enzyme.

## 7.2 The ω-TAm screen: an overview of the results and indications for candidate enzymes

As outlined in section 4.2.1, the ability to purchase aldehydes is severely limited due to the very reactive chemistries of the compounds. The first challenge was therefore to source 6-oxohexanoate or to find alternatives. The methyl-ester of 6-oxohexanoate is available to purchase and since the enzymes of the pathway are cloned individually there is the added benefit of using these to biocatalytically produce 6-oxohexanoate. In addition the methyl-ester could also be hydrolysed to yield 6-oxohexanoate and stored at -20$^{o}$C until required. All three were tested and although in the ester, the methyl group is at the opposite end to the molecule to the hydroxyl group and therefore steric hindrance should be minimal, the comparison between the ester and the biocatalytically produced 6-oxohexanoate was more severe than expected. In 12 of the 23 TAm's initially tested, the ranking scores obtained differed by more than 8 places and in some cases the difference was even more significant. This pointed towards the model methyl ester substrate being an unsuitable substitute for 6-oxohexanoate due to too large a discrepancy between the two. More information on the TAm structure and docking studies would need to be completed to fully explain why this is so.

A total of 28 TAm's were screened either on chemically or biocatalytically produced 6-oxohexanoate in a design which again used spectrophotometry and followed co-factor changes. After the production of pyruvate by the transamination of Alanine, LDH was used to convert this into lactate with the corresponding absorbsion change as NADH is converted to $NAD^{+}$. The first TAm screen was completed using biocatalytically produced 6-oxohexanoate and thus, as explained in 4.2.2, the overall absorbance increased since there was still likely copious amounts of caprolactone (the maximum produced in the initial reaction was 0.625 mM based on a 1:1 ratio of caprolactone to $NAD^{+}$ leaving approximately

43.375 mM of unreacted caprolactone) and likely 6-hydroxyhexanoate present. Thus the $NAD^+$ released from the LDH reaction drove the dehydrogenation reaction to produce more 6-oxohexanoate. Because of the rate of reaction difference between the TAm reaction and the dehdyrogenation, this led to a net increase in absorbance. As expected in the next two screens where the chemically derived 6-oxohexanoate was used, the absorbance decreased since there were no enzymes and other substrates present to convert the $NAD^+$ back into NADH. At the end of the three screening experiments the TAms were ranked by taking into account the largest absorbance change on 6-oxohexanoate and the smallest on cyclohexanone. Looking at the three screening experiments there were at times several experiments where the results were drastically different to the other two. When this was the case the outlier was not taken into account e.g. with pQR810, experiment 1 and 2 were in close agreement however 3 was vastly different and so wasn't included. Possible reasons for this difference could be attributed to experimental error, for example if the initial reaction rate was large and then quickly slowed, but the plate was slower to be read than in another experiment (e.g. slight delays in putting the plate in the reader or pipetting was slower than previously) then the result could look as if the clone didn't produce a large change relative to time 0 when previously it had done. It is interesting that only 7 from the cyclohexanone screen have one number not taken into account yet the 6-oxohexanoate reactions have 14. This could show the sensitive nature of working with an aldehyde. Moreover a total of 7 from the biocatalytically produced method were not included compared to 4 and 3 in experiments 2 and 3 when using chemically synthesised 6-oxohexanoate. A possible reason for this is that cell lysates were used and even though they were very dilute, aldehyde dehydrogenases from *E.coli* BL21 (DE3) will be present and possibly some variability was introduced here. Also, since the maximum amount of 6-oxohexanoate produced enzymatically is 0.625 mM (assuming all 6-hydroxyhexanoic acid had been converted to 6-oxohexanoate) certain enzymes may have different kinetic

properties and behave differently with the lower concentration of 6-oxohexanoate from the enzymatic route as compared to a higher concentration from the chemically pure method (working concentration of 1 mM). Due to these errors the screen was designed to only give an indication of likely candidates and therefore the top three ranked were chosen to be cloned. These were pQR810 (*Psuedomonas putida* PP5182), pQR983 (*Deinococcus geothermalis* DGEO_2743) and pQR1014 (*Micrococcus luteus* MLUT_00920).

Both PP5182 and DGEO_2743 were cloned using CPEC into both operon versions, which interestingly only worked using lower cycle numbers but also a higher temperature than a reaction with more DNA parts. The reason for this is unclear since it is an assembly reaction and not an amplification. A *BamHI* site was cloned between the RBS and the TAm genes to act as a spacer as well as a restriction site to confirm successful assembly. MLUT_00920 was significantly more difficult to clone and when assembled by CPEC, effectively knocked out all other genes in the operon. This was likely due to the high GC content (72%) of the gene which would make it more likely to bind to non-complimentary sites. Instead MLUT_00920 was cloned using restriction cloning by introducing *AvrII* and BbvCI sites which proved much more successful. It was only cloned into operon A however because GC-MS results indicated its poor performance (see later sections).

## 7.3 The problem with 6-ACA detection

Intermediate detection and quantification of the metabolites is arguably the most challenging part of the project. As outlined in section 5.1, amines require derivitisation for detection and in the case of GC-MS, to improve sensitivity and peak resolution by making the compound more volatile. An important observation that isn't necessarily immediately clear is the method by which is used to create a standard curve of 6-ACA on GC-MS. This is very important and crucial to achieving accurate results. After finding a small peak of 6-ACA present in pQR1556 samples grown on 200 mM cyclohexanol, 99% pure 6-ACA was used to

attempt to create a standard curve which presented several challenges. The only solvent found to dissolve 6-ACA directly was water which cannot be put down a GC silica column. Instead a concentrated solution of 10g/L was made and diluted into a set of standards. After drying and derivitising with BSTFA and TMCS, a large variety of peaks were observed at all concentrations from 160 mg/L to 20 mg/L with no single peak corresponding to 6-ACA as observed in samples taken from cultures. The question of how this is possible given that 6-ACA produced from the clones is effectively dissolved in a water base from the culture media is difficult to answer at first. The answer likely lies by looking at the equation of the condensation reaction of 6-ACA to produce an oligomer (Figure 7.1).

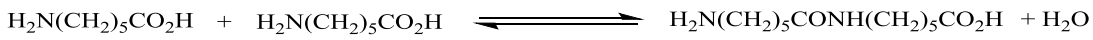



**Figure 7.1 The equation of reacting two 6-ACA monomers to create a oligomer releasing water**

The reason why 10 g/L or even 1 g/L 6-ACA was made up and different concentrations were diluted from this was because weighing the masses required to produce 1 ml of, for example 100 mg/L, were too small to handle so practically this makes sense. However, by looking at the equation to form an oligomer it can be reasoned that if the concentration of water becomes too low (i.e. a high concentration of 6-ACA is present) then the reversible reaction will be pushed to the right in order to increase the concentration of water by releasing it from a condensation reaction. Thus a solution of 10 g/L to 1 g/L would form oligomers which would be carried over to the standard dilutions. Presumably this would create a variety of oligomers and possibly even large ring molecules which explains why there are a variety of peaks detected in what is a 99% pure standard of 6-ACA (Figure 6.7). In fact, a paper from 1978 found that significant oligomers of 6-ACA can form at low

temperatures in water and that a high concentration of water (i.e. dilute 6-ACA solution) favours the monomers (Mares and Sheehan, 1978). This paper was also significant in understanding why, in this case, it is likely the oligomers and not caprolactam which is formed. They found that temperatures as high as 300$^{o}$C are needed to cyclise 6-ACA to form caprolactam and that even then oligomers were still found in water. They also found that approximately 19% of 6-ACA would cyclise into caprolactam at temperatures as low as 100$^{o}$C over a 6 hour period. This could explain why there is a difference of 50.84 mg/L of 6-ACA between drying at 75$^{o}$C for 16 hours or speed vac drying at room temperature. The higher the temperature, the more likely some degree will cyclise into caprolactam. The effect can also be seen more dramatically with 6-hydroxyhexanoic acid where there is a large increase of 319.1 mg/L of 6-hydroxyhexanoate which could possibly be down to the reaction of the alcohol and carboxylic acid groups cyclising to form caprolactone. It is likely therefore that if, for example several 100ml solutions of concentrations of 6-ACA are made at <160mg/L, the water concentration would be large enough to favour monomers yet the amounts could practically be weighed out. If then dried and derivatised less oligomers would be seen. However it is more complicated than that. Another problem was found when using only 99% pure 6-ACA and derivatising in ethyl acetate with BSTFA and TMCS. No water was present and 6-ACA doesn't dissolve directly into Ethyl acetate so stays in a solid phase and needed heating to 70$^{o}$C for 1 hour to derivatise. Therefore no oligomers are able to form. It was found that there is likely a significant degree of Acyl transfer from the Ethyl acetate upon heating. However when a set of 50 ml cultures of non induced pQR1553 cells at the end of fermentation were spiked with 6-ACA and processed in the same way as if the clones produced it. A very clear linear standard curve was produced. This Acyl transfer is not seen when samples or standards are taken from culture medium and this is likely attributed to more reactive amino acids in the media which are Acylated in lieu of 6-ACA. However, dilute 6-ACA standards were not dried from pure water and so it

can't be concluded whether this acyl transfer will only occur at high 6-ACA. In summary, 6-ACA must be in a dilute concentration in media and then dried and derivatised in order to create a linear standard curve.

## 7.4 A high producing clone is identified

The clones created based on indications from the TAm screen were initially grown on 10 mM cyclohexanol and their metabolic pathway profile was quantified. There are two important findings from this screen. Firstly, in all clones, the ones where the *Acinetobacter calcoaceticus* NCIMB 9871 genes were under the regulation of their natural RBS produced the most 6-ACA overall. In some cases the difference was significant such as between pQR1553 and pQR1554, the former was the overall best producer with a total of 138.70 mg/L under those drying conditions whilst the latter was only 53.18 mg/L; a considerable increase of 85.52 mg/L given the only difference between the two clones is the RBS. This could indicate that the subtle differences in the RBS may have large effects *in vivo*. This was the case with *ChnB*, in pQR1553 it's RBS is only GGAGA at position -6 whereas in pQR1554 it was AGGAG at position -7 yet pQR1554 had 76.62 mg/L more cyclohexanone than pQR1553. In that instance no caprolactone was detected so a low functioning *ChnC* cannot account for this difference i.e. a build up of caprolactone could substrate inhibit cyclohexanone monoxygenase therefore creating a build up of cyclohexanone. Secondly, the prediction that the rank order of TAm's should be *Psudomonas putida (PP5182), Deinococcus geothermalis (DGEO_2743)* followed by *Micrococcus luteus* (MLUT_00920) was slightly inaccurate. In fact pQR1557 produced 72.51 mg/L more 6-ACA than pQR1555 whilst the screen suggested that it would be closer between the two enzymes. However cyclohexanone concentration was important and pQR1553 had the lowest amount present at the end of fermentation.

After the drying method was shown to be very important to the amount of 6-ACA detected, pQR1553 was still the highest producing clone. Four different media were screened and showed that the medium drastically influenced the amount of 6-ACA produced from 10 mM cyclohexanol. The four media screened were BHI, 2xTY, NB and NZCYM. The make-up of these media is complex and have Nitrogen and Carbon supplies from different sources. NB performed the worst producing zero 6-ACA whilst 2xTY produced the largest amount achieving 264.48 mg/L from 10 mM cyclohexanol. This could suggest that the Nitrogen source is a critical factor. Both contain Yeast extract to provide a carbon source and vitamins plus some peptides however the largest source of Nitrogen is provided by either peptone or tryptone. Since NB and BHI contain peptone whilst NZCYM and 2xTY contain hydrolysed or tryptic digests of casein with the former two producing the lowest and the latter two producing the largest concentration of 6-ACA, the Nitrogen source appears to be important. Casein is a protein sourced from milk whilst peptone can be sourced by a variety of means ranging from soymilk to meat and so the Nitrogen content is variable. This could impact the amount of 6-ACA produced.

The substrate concentration of cyclohexanol also play a pivotal role in the production of 6-ACA. Although *in vitro* enzyme studies suggested that the majority of enzymes would remain active at high substrate concentrations (as high as 500 mM in some cases), the added complication of mass transfer as well as co-factor availability will impact the enzymes ability to replicate these results. Also the enzymes *in vivo* are at a significantly lower concentration than was tested *in vitro* which would also have an impact on the cyclohexanol concentration that could be utilised. The results suggested that after screening on three different concentrations of cyclohexanol, that at 30 mM, the largest concentration of 6-ACA was produced (222.3 mg/L) and at both 20 mM and 30 mM significantly higher levels of other intermediates were also present. However at 50 mM, the pathway really stalled at cyclohexanone. There was approximately 50:50 of cyclohexanol

and cyclohexanone but no other intermediates were detected. This is significant since it would appear that the overall total molar amounts of intermediates was also reduced with only 43.9% recovered compared to 76.8% accountable from 20 mM cyclohexanol. This could mean that over the course of the incubation period, because cyclohexanol could be slow to enter the cells because of the reduced solubility at this concentration, it partially evaporated. Also because there was no caprolactone present, the BVMO was therefore inactive and given that there was also substantial amounts of cyclohexanol present, it seems logical that the cyclohexanol dehydrogenase was substrate inhibited. It is also possible that at 500 mM and 37$^{o}$C, cyclohexanol is only partially soluble in water since at 30$^{o}$C it's soluble to 2.15 g in 50ml of water (source: ncbi database, 500 mM 2.5 g in 50ml water) so it is possible at the higher temperature it may not be available to the cells.

## 7.5 Addition of a second copy of *ChnD* and a T7 promoter and terminator offers no improvement

A second copy of *ChnD* under the control of its own T7 promoter and terminator was cloned in order to try and reduce the amount of 6-hydroxyhexanoate building up so as to hopefully convert more into 6-oxohexanoate for transamination. Whilst a marginal improvement in 6-ACA was achieved on 10 mM cyclohexanol in LB media compared to pQR1553, pQR1559 also had higher concentrations of adipic acid and 6-hydroxyhexanoate acid. This suggests that the transamination was a bottleneck since a build up of 6-oxohexanoate could be converted to adipic acid and also lead to a build up of 6-hydrohexanoate. Conversely, pQR1559 overall performed less well at higher substrate concentrations and on the different media. Whilst pQR1553 produced 222.3 mg/L of 6-ACA on 30 mM cyclohexanol in LB media, pQR1559 only produced 151.2 mg/L. Less 6-hydroxyhexanoate was also detected as well as adipic acid than the pQR1553

experiments. This data suggests a conflict with the 10 mM cyclohexanol data for pQR1559 because if there was a bottleneck in the transamination at these higher substrate concentrations then both adipic acid and 6-hydroxyhexanoate should have increased. Because solvent extraction wasn't done on the pQR1559 at higher cyclohexanol concentrations, the amounts of cyclohexanol, cyclohexanone and caprolactone can't be determined. It is entirely possible that at these higher concentrations there was a more significant build up of cyclohexanone and cyclohexanol in pQR1559. This could be feasible since there are indications of back end substrate inhibitions in pQR1559 on 10 mM which could also lead to a front end build up of metabolites which become more apparent at higher cyclohexanol substrate concentrations. Kinetic modelling of each enzyme step would need to be done in order to determine if the TAm step is indeed the rate limiting step and rate equatons would need to be calculated. In conclusion, none of the conditions explored on pQR1559 increased the concentration of 6-ACA achieved with pQR1553 on 10 mM cyclohexanol in 2xTY media.

# Chapter 8: Conclusion and future work

# 8 Conclusion

## 8.1 Restatement of aims and achievements

The aims of the project were initially split into three distinct phases. The first of these phases was to assemble an operon consisting of only the *Acinetobacter calcoaceticus* NCIMB 9871 genes. It was decided that it would be interesting to see what effect the RBS had on the intermediate profile and whether natural RBSS would be more productive since these may have evolved to finely tune the pathway flux by way of moderating enzyme expression levels. Two versions were assembled, one where all had their natural RBS and the other were they all had the same. Genomic samples of the *Acinetobacter calcoaceticus* NCIMB 9871 was used to amplify the four required genes. These genes exist as a cluster, directed in different orientations which made extracting the group together not an option. Instead this allowed the omittance of any gaps between the genes as well as it being easier to redesign the RBSS of each gene by way of simply altering the primer for the PCR. This phase of the project was anticipated to be the most challenging since assembling multiple genes in a single operon and plasmid has traditionally been difficult to do with restriction enzymes. Multiple unique restriction sites as well as several possible buffers would likely be needed. Coupled with potential problems which are inherent with PCR such as failing and the difficulty of assembling several DNA fragments (i.e. in terms of successful assembly with minimal false positives) a significant amount of time was thought to be needed. Instead however, because a nested approach was taken to each gene as well as successful PCR's, both operon versions were assembled very rapidly. Gibson assembly was initially attempted however this was unsuccessful and results suggested that only two of the four genes joined properly. Instead circular polymerase extension cloning (CPEC) was tried and worked very well. Throughout this work results varied in terms of how many successfully assembled plasmids were generated in each CPEC reaction although broadly speaking

around half of all colonies screened digested as expected. It's not quite clear why Gibson assembly was unsuccessful and would only join certain fragments up. The optimum conditions were not investigated however it did appear that cycle number had an affect; later on in the project when the TAm's were added into each operon they required a, fewer cycles to work as well as high gene GC content being problematic (see chapter 4.2.6).

The second phase of the project was to screen a library of ω-TAm's for the best activity on 6-oxohexanoate whilst having minimal on cyclohexanone. One screening method developed for testing for TAm activity which is popular in our research group involves using sMBA (S-Methylbenzylamine) as the amine donor and detecting the appearance of acetophenone at 254nm via HPLC. The larger the peak area for acetophenone, the greater the conversion of sMBA and therefore the better the enzyme activity. Instead of this, Alanine was used since *in vivo* it is one of the most abundant amino acids. The pyruvate generated after transamination was detected via an LDH assay. Overall this method is likely to have a degree of inaccuracy for a number of reasons. Firstly there were significant differences seen when using 6-oxohexanoate produced biocatalytically or chemically however this could be attributed to reduced amounts of 6-oxohexanoate theoretically present from the biocatalytically produced route. However three candidates were found to exhibit good activity on 6-oxohexanoate whilst having minimal on cyclohexanone and these were PP5182 from *Pseudomonas putida,* DGEO_2743 from *Deinococcus geothermalis* and MLUT_00920 from *Micrococcus luteus* and were ranked in that order. Once pathway intermediates were detected and quantified by GC-MS, the actual order of best producing TAm was PP5182, MLUT_00920 and DGEO_2743 in the same operon construct with the natural RBS sites for the *Acinetobacter calcoaceticus* genes. Although the actual rank was slightly different to what was predicted, the important finding was that taking into account the activity on cyclohexanone was very important. *PP5182* was not identified as having the largest activity on 6-oxohexanoate by a considerable margin, however it did have the

lowest activity on cyclohexanone. Importantly all three showed very little cyclohexylamine formation in either operon version which also helps to validate the TAm screening LDH assay. Looking at figure 34 in chapter 4.2.4 one can see that several TAms had either very good activity on either molecule or very poor activity. It is also likely that pQR804 might have been a potential target for incorporation ranking 4$^{th}$ overall and was the only other enzyme to exhibit good activity on 6-oxohexanoate and a low amount on cyclohexanone. Importantly 15 of the 28 TAm's showed very good activity on cyclohexanone also.

The final phase of the project came after cloning in the ω-TAms into each operon (with the exception of *MLUT_00920* into operon B; artificial RBS); the intermediate profile of cells grown with media supplemented with cyclohexanol was then identified. The first three molecules in the pathway have markedly different chemistries to the latter half in terms of boiling points and solubility in solvents. This made detecting these by GC-MS simple and by mixing Ethyl acetate with the sonicated culture samples and centrifuging, the molecules pulled easily out of solution into the organic layer. 6-hydroxyhexanoate, 6-ACA, adipic acid and cyclohexylamine however have higher boiling points and were not very soluble in the Ethyl acetate. In this instance the samples were dried and derivatised in a molar excess of BSTFA and TMCS to increase volatility and solubility in the organic solvent. An important find was that the method of preparing the standards of 6-ACA was important. If it was prepared in a concentrated solution ( concentrations less than 1 g/L were not tested) in Ethyl acetate, derivatised and diluted at various concentrations, significant amounts of acyl-transfer appeared to have occurred. Alternatively if it was diluted from a concentrated solution in water (the only solvent to dissolve 6-ACA directly) and subsequently dried, then significant polymerisation appeared to be initiated. This is thought to be caused by the equilibrium moving to the right and releasing water to counteract the low concentration of water in a high concentration of 6-ACA solution (>1g/L only tested). To circumvent this, a standard curve was made by making different concentrations of 6-ACA in non-induced cell

cultures grown in LB and processed in the same way. That way alternative amino acids were thought to mop up the acyl transfer. Also because the volume of the culture was larger, the smaller concentrations of 6-ACA were easier to weigh. this method produced a linear standard curve. In future standards must be prepared like this or possibly in low concentrations and at high volumes of water e.g. 500 ml at 100 mg/L.

Although it turned out that the method of drying the samples impacted on the amount of 6-ACA that was detected (i.e. speed vac compared to leaving to evaporate), the highest producing clone on 10 mM cyclohexanol was pQR1553 i.e. natural RBSS for the *Acinetobacter caloaceticus* genes along with *PP5182*. This validated that the ω-TAm screen was correct as well as highlighting the large effect the RBS has on the intermediate profile. Although when expressed, by eye, there appeared to be no or very little change in enzyme concentration between the two operon versions, there must've been some at a level to cause significant differences. In all cases, operon A was the highest producer of 6-ACA and in the case of *PP5182,* a difference of 85.52 mg/L. From experiments done on the individual enzymes from *Acinetobacter* as well as information from the literature for the CHMO, it was highlighted that there would likely be large amounts of cyclohexanone and 6-hydroxyhexanoate, the former because of substrate and product inhibition, the latter due to poor activity. After looking at the intermediate profile of pQR1553 in more depth these indications, at 10 mM were partially correct in that only very small amounts of cyclohexanone were present however 6-hydroxyhexanoate was present at the highest concentration. When grown on 10 mM cyclohexanol in LB media pQR1553 achieved an average of 189.54 mg/L 6-ACA. The choice of media actually had a large effect on the 6-ACA levels achieved as well as the intermediate profile in general. In fact 2xTY media was the best where cultures grown on 10 mM cyclohexanol produced 264.48 mg/L (2.02 mM). A cyclohexanol concentration screen was done using 20 mM, 30 mM and 50 mM cyclohexanol in LB media. 30 mM produced the most amount of 6-ACA at 223.3 mg/L

representing a 0.26 mM increase from 10 mM cyclohexanol. Interestingly however cyclohexanone and 6-hydroxyhexanoate concentrations rose sharply (an 8.67 mM and 2.49 mM increase respectively) and even at 20 mM cyclohexanone was found to be at over 845.82 mg/L, a sharp increase from 10 mM cyclohexanol. Adipic acid concentrations also increased with higher concentrations of cyclohexanol with the exception of 50 mM where no compounds after cyclohexanone were identified. This increase could be attributed to a build up of 6-oxohexanoate for which the TAm's couldn't react with fast enough leading to high activity from aldehyde dehydrogenases.

Whereas cyclohexanone levels seem to depend on the concentration of cyclohexanol fed in bolus, 6-hydroxyhexanoate was consistently shown to be present at high levels in all cyclohexanol concentrations and media screens. Efforts to reduce this by cloning in a second, independently expressed *ChnD* into the plasmid increased 6-ACA levels from 10 mM cyclohexanol but reduced the amounts seen at higher substrate concentrations as well as achieving lower titres when grown on the various different media. Although no solvent extraction phase were ran for these experiments, there is a potential that the ω-TAm (*PP5182*) couldn't process the excess 6-oxohexanoate quick enough which led to an increase in 6-hydroxyhexanoate and adipic acid. More investigation into this would need to be completed to determine the metabolic profile better however the bottom line is that at no conditions tested (i.e. the same ones for pQR1553) did pQR1559 outperform pQR1553 in terms of 6-ACA production. In conclusion the largest amount of 6-ACA observed at 264.48 mg/L was when pQR1553 was grown on 2xTY media and then supplemented with 10 mM cyclohexanol.

## 8.2 Yield comparison from published results

In order to assess the success of the 6-ACA titre, it is important to, where possible, compare these results with published results from the literature. Whilst it is known that there are several industrial chemical companies which are manufacturing 6-ACA biologically; yield information from these is limited at best. Even still, published research for producing 6-ACA is also scarce. Zhou *et al* published results using design of experiments and in a 32 member fractional factorial library which simultaneously compared the expression levels and media composition and found optimum conditions. They have engineered an *E.coli* strain to produce 6-ACA from α-ketoglutarate using 6 heterologous enzymes each under the control of their own promoter (Zhou *et al.*, 2015). They claim that in shake flask culture they increased titres from 9mg/L to 48mg/L and when grown in a 10L fermenter less side products were found and a titre of 160mg/L was achieved. Their work was interesting as it showed the power of a combined DoE method and the importance of the correct promoter and RBS as well as an optimum media. A different approach was taken by Sattler *et al* who used a combination of enzymes to convert cyclohexanol into 6-ACA in very similar to the work presented in this thesis with several notable differences (Sattler *et al.*, 2014). Firstly their work was wholly *in vitro* and used purified enzymes in a one pot reaction. Secondly they considered 6-hydroxyhexanoic acid a dead end product and used methanol as a nucleophile to open caprolactone which would cap the carboxyl group. The reason for this is they had initially tested a library of primary alcohol dehydrogenases, including the 6-hydroxyhexanoate dehydrogenase from *Acinetobacter calcoaceticus* NCIMB 9871 that is presented in this thesis however this was one of only two that worked yet they had problems expressing *ChnD* and noted its inability to tolerate high substrate loading; this reaffirms the findings in this thesis. They found that, rather surprisingly, when the carboxylic acid moiety was converted into a methyl ester (i.e. ring openning of caprolactone with methanol), that the activity of all the alcohol

dehydrogenases greatly increased which suggested the actual carboxyl group was inhibitory to the oxidation. They then used an alcohol dehydrogenase to create 6-oxohexanoic acid and then subsequently an ω-TAm to create the 6-ACA. Their yields were 12mM 6-ACA (1.57g/L) from a 50mM cyclohexanol substrate concentration and after 20h 17.5mM cyclohexanol was remaining. The supplementary data shows that each enzyme was purified and used in milligram quantities. Although the work was impressive from a yield point of view, industrial use would be limited due to the massive cost of the enzymes and co-factor which would be incurred which would ultimately prove inhibitory to large scale manufacture.

The work presented by Turk *et al* shows they cloned 11 enzymes onto plasmids and utilised TCA intermediates for substrates to produce 6-ACA. They achieved yields of 160 mg/L 6-ACA from batch shake flask experiments. Of the three papers described above, it can be seen that the results from the work described in this thesis show a marked larger titres in comparison with other *in vivo* processes. With very little in depth optimisation, titres were still higher than even those achieved after DoE as well as from concentrating the cells down (Turk *et al.*, 2015).

Although titres achieved were relatively high when compared with research producing 6-ACA, in a wider context it still has some way to go to catch-up with amounts produced of other biopolymer monomers. Kind *et al* have produced a clone of *C.glutamicum* which firstly over-produced Lysine from glucose and then further engineered it to express a codon-optimised Lysine decarboxylase to produce the diamine cadaverine (aka diaminopentane) and achieved titres of 88g/L from fed batch fermentations (Kind *et al.*, 2014). The fact that the genes were incorporated into the genome as well as the fact that Lysine is a naturally occurring amino acid which is then decarboxylated are likely to have contributed to achieving this titre. Another Nylon precursor which is also produced in the

pathway described in this project is adipic acid. Of great interest in the bioproduction of adipic acid is the precursor *cis cis* muconic acid which chemically hydrogenated to produce adipic acid. Although usually produced from benzene, Niu *et al,* produced it from D-glucose in *E.coli* cells by transforming 6 enzymes and achieved an overall titre of 36.8 g/L from a fed batch fermentation (Niu, Draths and Frost, 2002). An interesting alternative to *cis cis* muconic acid production was by Jung *et al* who used *Klebsiella pneumoniae* which naturally expresses all enzymes to produce ccMA however lacks the necessary carbon flux to produce it. The group engineered the strain to increase the amount of catechol (a ccMA precursor) and subsequently over express the *catA* gene (catechol 1,2-dioxygenase) which converts it to ccMA. Further deletions of the degradative enzymes media improvements achieved a titre of 2.1g ccMA (Jung, Jung and Oh, 2015). Putrescine (1,4-diaminobutane), a 4 carbon diamine is produced naturally in *E.coli* and is combined with adipic acid to form polyamide 6,4. This is currently produced at around 10,000 tonnes per annum and so is another molecule with potential interest. Qian *et al,* have engineered an *E.coli* strain to produce Putrescine by deleting degradation pathways and achieved a titre of 1.68g/L with a total conversion of 0.168 g/g of glucose (in comparison, the pathway described in this thesis produced $0.268 g_{6\text{-}ACA}/g_{cyclohexanol}$). After fermentation and high cell culture a maximum of 24.2 g/L was achieved (Qian, Xia and Lee, 2009).

Finally, GABA (γ-aminobutanoic acid) is a 4 carbon homologue of 6-ACA and has been produced from the decarboxylation of glutamate by Le Vo *et al* and achieved a total titre of 5.5 g/L (Vo, Kim and Hong, 2012).

## 8.3 Future work

### 8.3.1 Design of Experiments and fermentation

The original aims set out for this project have been completed however there are several avenues with which to progress this project. One option with regards to the current pQR1553 plasmid is to fully optimise the feeding strategy and growth conditions. The work presented in this thesis outline a potential starting point and since much time has been in developing the analytics and overcoming associated technical problems, very little time was left to explore the optimum conditions. Despite this, yields are comparatively high to what is currently published in the literature and by utilising DoE, factor interactions, ideal promoter and RBS designs as well as optimum conditions can be found. As explained above, there is a theory that the source of protein is important and so the make-up of the media and could also be investigated and an ideal minimal media could be designed. Potential factors to investigate include the rate of feeding of cyclohexanol (fed-batch or continuous), stirrer speed in a fermenter, pH and temperature. Whether a spike of cyclohexanol, or indeed the rate at which the 10 mM cyclohexanol was metabolised, or a continual dilute feed would be most ideal would need to be investigated. This work showed that high concentrations of cyclohexanol supplied in a bulk volume is detrimental to 6-ACA production however, can it be said that a 200 mM solution fed slowly over several hours is detrimental? I don't believe so. I feel it is likely that the cyclohexanol is metabolised quickly however the evolution of 6-ACA also needs to be identified.

### 8.3.2 Coupling with central metabolism and further plasmid engineering

In the effort to produce 6-ACA in a wholly green method of manufacture, ultimately the synthesis of cyclohexanol will need to be coupled to the bacteria's central metabolism. Currently there is a PhD project in the Ward lab which is undertaking this. Schikimic acid is a molecule produced from glucose in the central metabolism of bacteria in the common aromatic biosynthetic pathway and is responsible for several amino acid synthesis (Herrmann, 1995). A gene cluster from *Streptomyces rishiriensis* (*MycA4-1* ) is being integrated into the pathway and subsequent thioesterases, decarboxylases and P450's are being investigated to generate the cyclohexanol in a pathway shown in figure 1. Owing to the potential size of such a plasmid as well as taking into account the five genes in the operon from this work, genome integration is likely to be required to reduce the cellular burden.

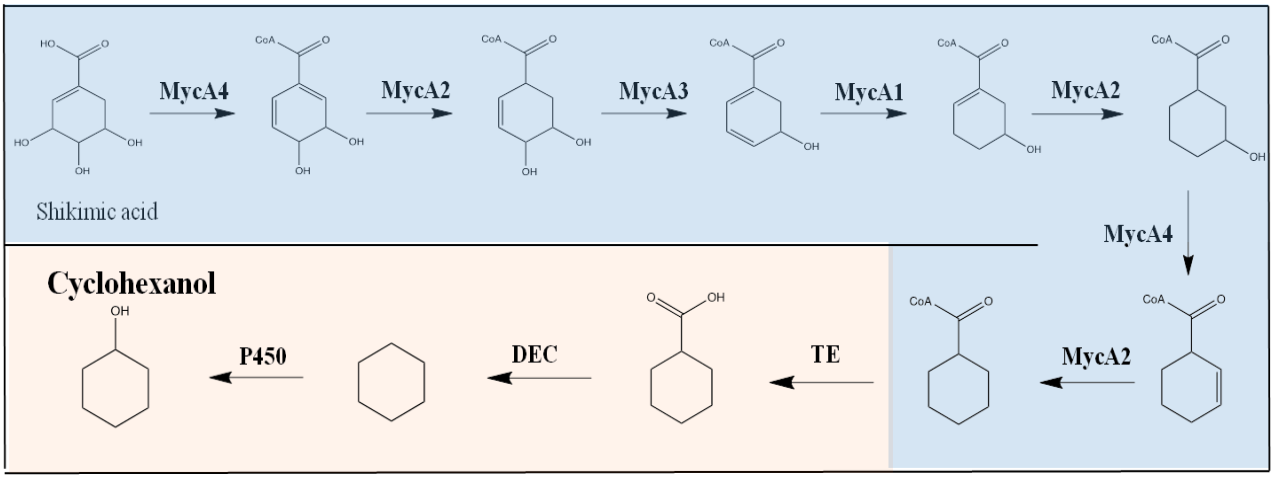



**Figure 8.1 Proposed pathway for cyclohexanol production from central metabolism**

Aside from integrating with the genome, future work could be carried out on the pQR1553 plasmid. The RBS has been shown to have a dramatic effect on the production titre of 6-ACA and the intermediate profile and so further mutations and optimisations of this by introducing point changes and deletions could also be explored.

# Bibliography


Abubackar, H. N., Veiga, M. C. and Kennes, C. (2012) 'Biological conversion of carbon monoxide to ethanol: Effect of pH, gas pressure, reducing agent and yeast extract', *Bioresource Technology*. Elsevier Ltd, 114, pp. 518–522. doi: 10.1016/j.biortech.2012.03.027.

Abubackar, H., Veiga, M. and Kennes, C. (2015) 'Ethanol and Acetic Acid Production from Carbon Monoxide in a Clostridium Strain in Batch and Continuous Gas-Fed Bioreactors', *International Journal of Environmental Research and Public Health*, 12, pp. 1029–1043. doi: 10.3390/ijerph120101029.

Akhtar, M. K., Turner, N. J. and Jones, P. R. (2013) 'Carboxylic acid reductase is a versatile enzyme for the conversion of fatty acids into fuels and chemical commodities.', *Proceedings of the National Academy of Sciences of the United States of America*, 110(1), pp. 87–92. doi: 10.1073/pnas.1216516110.

Akita, H., Nakashima, N. and Hoshino, T. (2016) 'Pyruvate production using engineered Escherichia coli', *AMB Express*. Springer Berlin Heidelberg, 6(94), pp. 1–8. doi: 10.1186/s13568-016-0259-z.

Alphand, V. *et al.* (2003) 'Towards large-scale synthetic applications of Baeyer-Villiger monooxygenases', *Trends in Biotechnology*, 21(7), pp. 318–323. doi: 10.1016/S0167-7799(03)00144-6.

American Chemical Society (1995) *The First Nylon Plant*.

Antron (2013) 'The Difference between Type 6 , 6 and Type 6 Nylon', pp. 1–3.

Bart, J. C. J. and Cavallaro, S. (2015) 'Transiting from Adipic Acid to Bioadipic Acid. Part II. Biosynthetic Pathways', *Industrial & Engineering Chemistry Research*, 54, pp. 567–576. doi: 10.1021/ie502074d.

Beardslee, T. and Picataggio, S. (2012) 'Bio-based adipic acid from renewable oils', *Lipid Technology*, 24(10), pp. 223–225. doi: 10.1002/lite.201200230.

Behura, S. K. and Severson, D. W. (2013) 'Codon usage bias: Causative factors,

quantification methods and genome-wide patterns: With emphasis on insect genomes', *Biological Reviews*, 88, pp. 49–61. doi: 10.1111/j.1469-185X.2012.00242.x.

Bellussi, G. and Perego, C. (2000) 'Industrial catalytic aspects of the synthesis of monomers for nylon production', *Cattech*, 4(1), pp. 1–13.

BENNETT, A. P. *et al.* (1988) 'Purification and Properties of caprolactone Hydrolases from Acinetobacter NCIB 9871 and Nocardia globevula CLl', *Journal of general microbiology*, 134, pp. 161–168.

Bryksin, A. V. and Matsumura, I. (2010) 'Overlap extension PCR cloning: A simple and reliable way to create recombinant plasmids', *BioTechniques*, 48(6), pp. 463–465. doi: 10.2144/000113418.

Buschke, N. *et al.* (2013) 'Systems metabolic engineering of xylose-utilizing Corynebacterium glutamicum for production of 1,5-diaminopentane', *Biotechnology Journal*, 8(May), pp. 557–570. doi: 10.1002/biot.201200367.

Chen, Y. C., Peoples, O. P. and Walsh, C. T. (1988) 'Acinetobacter cyclohexanone monooxygenase: gene cloning and` sequence determination.', *Journal of Bacteriology*, 170(2), pp. 781–789.

Cheng, Q. *et al.* (2000) 'Genetic Analysis of a Gene Cluster for Cyclohexanol Oxidation in Acinetobacter sp. SE19 by in vitro transposition', *JOURNAL OF BACTERIOLOGY*, 182(17), pp. 4744–4751. doi: 10.1128/JB.182.17.4744-4751.2000.Updated.

Cramer, H., Verma, S. and Kiselewsky, M. (2011) *Derivatization and Separation of Aliphatic Amines*, *AnalytiX*. Available at: http://www.sigmaaldrich.com/technical-documents/articles/analytix/derivatization-aliphatic-amines.html (Accessed: 7 September 2017).

Dangel, W. *et al.* (1989) 'Enzyme reactions involved in anaerobic cyclohexanol metabolism by a denitrifying', *Archiives of microbiology*, pp. 273–279.

Donoghue, N. a and Trudgill, P. W. (1975) 'The metabolism of cyclohexanol by Acinetobacter NCIB 9871.', *European journal of biochemistry / FEBS*, 60, pp. 1–7. doi: 10.1111/j.1432-1033.1975.tb20968.x.

Dros, A. B. *et al.* (2015) 'Hexamethylenediamine (HMDA) from fossil- vs. bio-based routes: an economic and life cycle assessment comparative study', *Green Chem.* Royal Society of

Chemistry, 17(10), pp. 4760–4772. doi: 10.1039/C5GC01549A.

epa.gov (2017) *Nitrous Oxide Emissions*. Available at: https://www.epa.gov/ghgemissions/global-greenhouse-gas-emissions-data (Accessed: 7 September 2017).

Gibson, D. G. *et al.* (2009) 'Enzymatic assembly of DNA molecules up to several hundred kilobases', *Nature methods*, 6(5), pp. 12–16. doi: 10.1038/NMETH.1318.

Hanai, T., Atsumi, S. and Liao, J. C. (2007) 'Engineered synthetic pathway for isopropanol production in Escherichia coli', *Applied and Environmental Microbiology*, 73(24), pp. 7814–7818. doi: 10.1128/AEM.01140-07.

Heintz, S. *et al.* (2017) 'Development of In Situ Product Removal Strategies in Biocatalysis Applying Scaled-Down Unit Operations', *Biotechnology and bioengineering*, 114(3), pp. 600–609. doi: 10.1002/bit.26191.

Herrmann, K. M. (1995) 'The Shikimate Pathway as an Entry to Aromatic Secondary Metabolism'', *Plat Physiol*, 107, pp. 7–12.

Ho Hong, S. and Yup lee, S. (2004) 'Enhanced Production of Succinic Acid by Metabolically Engineered', *Biotechnology and Bioprocess Engineering*, 9, pp. 252–255.

Jung, H.-M., Jung, M.-Y. and Oh, M.-K. (2015) 'Metabolic engineering of Klebsiella pneumoniae for the production of cis,cis-muconic acid', *Applied Microbiology and Biotechnology*, 99(12), pp. 5217–5225.

Jung, Y. K. and Lee, S. Y. (2011) 'Efficient production of polylactic acid and its copolymers by metabolically engineered Escherichia coli', *Journal of Biotechnology*. Elsevier B.V., 151(1), pp. 94–101. doi: 10.1016/j.jbiotec.2010.11.009.

Kataoka, H. (1996) 'Derivatization reactions for the determination of amines by gas chromatography and their applications in environmental analysis', *Journal of chromatography A*, 733, pp. 19–34.

Kind, S. *et al.* (2014) 'From zero to hero - Production of bio-based nylon from renewable resources using engineered Corynebacterium glutamicum', *Metabolic Engineering*, 25, pp. 113–123. doi: 10.1016/j.ymben.2014.05.007.

Kunjapur, A. M. and Prather, K. L. J. (2015) 'Microbial engineering for aldehyde synthesis',

*Applied and Environmental Microbiology*, (January), p. AEM.03319-14. doi: 10.1128/AEM.03319-14.

Lee, J. W. *et al.* (2011) 'Microbial production of building block chemicals and polymers', *Current Opinion in Biotechnology*. Elsevier Ltd, 22(6), pp. 758–767. doi: 10.1016/j.copbio.2011.02.011.

Li, M. Z. and Elledge, S. J. (2007) 'Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC.', *Nature methods*, 4(3), pp. 251–256. doi: 10.1038/nmeth1010.

M.Clomburga, J. *et al.* (2015) 'Integrated engineering of β-oxidation reversal and ω-oxidation pathways for the synthesis of medium chain ω-functionalized carboxylic acids', *Metabolic Engineering*, 28, pp. 202–212.

Maddipati, P. *et al.* (2011) 'Ethanol production from syngas by Clostridium strain P11 using corn steep liquor as a nutrient replacement to yeast extract', *Bioresource Technology*. Elsevier Ltd, 102(11), pp. 6494–6501. doi: 10.1016/j.biortech.2011.03.047.

Mares, F. and Sheehan, D. (1978) 'Caprolactam Formation from 6-Aminocaproic Acid, Ester, and Amide', *Ind. Eng. Chem. Process Des, Dev*, 17(1), pp. 9–16.

Mathew, S. *et al.* (2013) 'High throughput screening methods for ω-Transaminases', *Biotechnology and Bioprocess Engineering*, 18, pp. 1–7. doi: 10.1007/s12257-012-0544-x.

McGrew, D. (2010) 'Direct route Fuel circle Getting to the point : Direct bio-b', *Specialty chemicals magazine*.

Mirza, I. A. *et al.* (2009) 'Crystal Structures of Cyclohexanone Monooxygenase Reveal Complex Domain Movements and a Sliding Cofactor', *JACS*, (15), pp. 8848–8854.

Na, D., Kim, T. Y. and Lee, S. Y. (2010) 'Construction and optimization of synthetic pathways in metabolic engineering', *Current Opinion in Microbiology*. Elsevier Ltd, 13(3), pp. 363–370. doi: 10.1016/j.mib.2010.02.004.

Nakamura, C. E. and Whited, G. M. (2003) 'Metabolic engineering for the microbial production of 1,3-propanediol', *Current Opinion in Biotechnology*, 14(5), pp. 454–459. doi: 10.1016/j.copbio.2003.08.005.

Niu, W., Draths, K. M. and Frost, J. W. (2002) 'Benzene-free synthesis of adipic acid',

*Biotechnology Progress*, 18(Cm), pp. 201–211. doi: 10.1021/bp010179x.

panap (2010) *Dumping of ammonium sulfate fertilizers is harmful to the Filipino people | PANAP*.

Park, E. and Shin, J. (2013) '□-Transaminase from Ochrobactrum anthropi Is Devoid of Substrate', *asm*, 79(13), pp. 4141–4144. doi: 10.1128/AEM.03811-12.

Park, S. J. *et al.* (2013) 'Synthesis of nylon 4 from gamma-aminobutyrate (GABA) produced by recombinant Escherichia coli', *Bioprocess and Biosystems Engineering*, 36, pp. 885–892. doi: 10.1007/s00449-012-0821-2.

Park, Y. S. *et al.* (2007) 'Design of 5'-untranslated region variants for tunable expression in Escherichia coli', *Biochemical and Biophysical Research Communications*, 356, pp. 136–141. doi: 10.1016/j.bbrc.2007.02.127.

Parker, L. (2009) *U. S. Global Climate Change Policy: Evolving Views on Cost, Competitiveness*, *Congressional research service*.

Poliakoff, M. *et al.* (2002) 'Green chemistry: science and politics of change.', *Science (New York, N.Y.)*, 297, pp. 807–810. doi: 10.1126/science.297.5582.807.

Pray, L. (2008) *Recombinant DNA Technology and Transgenic Animals*, *Nature Education*.

Qian, Z., Xia, X. and Lee, S. Y. (2009) 'Metabolic Engineering of Escherichia coli for the Production of Putrescine : A Four Carbon Diamine', *Biotechnology and bioengineering*, 104(4), pp. 651–662. doi: 10.1002/bit.22502.

Quan, J. and Tian, J. (2009) 'Circular Polymerase Extension Cloning of Complex Gene Libraries and Pathways', *PLoS ONE*, 4(7), p. e6441. doi: 10.1371/journal.pone.0006441.

Quan, J. and Tian, J. (2011) 'Circular polymerase extension cloning for high-throughput cloning of complex and combinatorial DNA libraries', *Nature protocols*. Nature Publishing Group, 6(2), pp. 242–251. doi: 10.1007/978-1-62703-764-8_8.

Ravishankara, A. R., Daniel, J. S. and Portmann, R. W. (2009) 'Nitrous Oxide (N2O): The Dominant Ozone-Depleting Substance Emitted in the 21st Century', *Science*, 326(5949), pp. 123–125.

Richter, N. *et al.* (2015) 'ω-Transaminases for the amination of functionalised cyclic ketones', *Organic & Biomolecular Chemistry*. Royal Society of Chemistry, 13, pp. 8843–

8851. doi: 10.1039/c5ob01204j.

Rosalam, S. and England, R. (2006) 'Review of xanthan gum production from unmodified starches by Xanthomonas comprestris sp.', *Enzyme and Microbial Technology*, 39, pp. 197–207. doi: 10.1016/j.enzmictec.2005.10.019.

Rudat, J., Brucher, B. R. and Syldatk, C. (2012) 'Transaminases for the synthesis of enantiopure beta-amino acids', *AMB Express*. Springer Open Ltd, 2(1), p. 11. doi: 10.1186/2191-0855-2-11.

Sattler, J. H. *et al.* (2014) 'Introducing an In Situ Capping Strategy in Systems Biocatalysis To Access 6-Aminohexanoic acid', *Angewandte Chemie International Edition*, 53, pp. 14153–14157. doi: 10.1002/anie.201409227.

Sayer, C. *et al.* (2014) 'The substrate specificity , enantioselectivity and structure of the ( R ) -selective amine : pyruvate Transaminase from Nectria haematococca', *FEBS*, 281, pp. 2240–2253. doi: 10.1111/febs.12778.

Schmidt, S. *et al.* (2015) 'An Enzyme Cascade Synthesis of ε-Caprolactone and its Oligomers', *Angewandte Chemie International Edition*, p. n/a-n/a. doi: 10.1002/anie.201410633.

Schneider, J. and Wendisch, V. F. (2010) 'Putrescine production by engineered Corynebacterium glutamicum', *Applied Microbiology and Biotechnology*, 88(4), pp. 859–868.

Sheng, D., Ballou, D. P. and Massey, V. (2001) 'Mechanistic Studies of Cyclohexanone Monooxygenase : Chemical Properties of Intermediates Involved in Catalysis †', *Biochemistry*, pp. 11156–11167.

Simon, R. C. *et al.* (2014) 'Recent Developments of Cascade Reactions Involving ω ‐ Transaminases', *ACS catalysis*, 4(1), pp. 129–143.

Sivashanmugam, A. *et al.* (2009) 'Practical protocols for production of very high yields of recombinant proteins using Escherichia coli', *Protein Science*, 18(1), pp. 936–948. doi: 10.1002/pro.102.

Speltz, E. B. and Regan, L. (2013) 'White and green screening with circular polymerase extension cloning for easy and reliable cloning', *Protein Science*, 22, pp. 859–864. doi:

10.1002/pro.2268.

The Essential Chemical Industry (2013) *Polyamides*.

Thomas, J. M. and Raja, R. (2005) 'Design of a "green" one-step catalytic production of epsilon-caprolactam (precursor of nylon-6).', *Proceedings of the National Academy of Sciences of the United States of America*, 102(39), pp. 13732–13736. doi: 10.1073/pnas.0506907102.

Turk, S. C. H. J. *et al.* (2015) 'Metabolic Engineering toward Sustainable Production of Nylon ‐ 6', *ACS synthetic biology*. doi: 10.1021/acssynbio.5b00129.

Usui, Y. and Sato, K. (2003) 'A green method of adipic acid synthesis: organic solvent- and halide-free oxidation of cycloalkanones with 30% hydrogen peroxide', *Green Chemistry*, 5, p. 373. doi: 10.1039/b305847f.

Vo, T. D. Le, Kim, T. W. and Hong, S. H. (2012) 'Effects of glutamate decarboxylase and gamma-aminobutyric acid (GABA) transporter on the bioconversion of GABA in engineered Escherichia coli', *Bioprocess Biosyst Eng.*, 35(4), pp. 645–650.

Ward, J. and Wohlgemuth, R. (2010) 'High-Yield Biocatalytic Amination Reactions in Organic Synthesis', *Current organic chemistry*, 1. doi: 10.2174/138527210792927546.

Wolfe, A. J. (2008) *Nylon: A Revolution in Textiles | Chemical Heritage Foundation*.

*World Nylon Fiber Report - Highlights* (2013). Available at: http://www.yarnsandfibers.com/preferredsupplier/spreports_fullstory.php?id=641 (Accessed: 7 September 2017).

Wu, H. *et al.* (2017) 'Characterization of Four New Distinct ω -Transaminases from Pseudomonas putida NBRC 14164 for Kinetic Resolution of Racemic Amines and Amino Alcohols', *Appl Biochem Biotechno*. Applied Biochemistry and Biotechnology, (79), pp. 972–985. doi: 10.1007/s12010-016-2263-9.

Yim, H. *et al.* (2011) 'Metabolic engineering of Escherichia coli for direct production of 1,4-butanediol.', *Nature chemical biology*. Nature Publishing Group, 7(7), pp. 445–452. doi: 10.1038/nchembio.580.

Yu, J.-L. *et al.* (2014) 'Direct biosynthesis of adipic acid from a synthetic pathway in recombinant Escherichia coli.', *Biotechnology and bioengineering*, 111(12), pp. 1–17. doi:

10.1002/bit.25293.

Zhou, H. *et al.* (2015) 'Algorithmic co-optimization of genetic constructs and growth conditions : application to 6-ACA , a potential nylon-6 precursor', *Nucleic acids researchs*, 43(21), pp. 10560–10570. doi: 10.1093/nar/gkv1071.

# Appendix

# Appendix A

This section contains the sequences and primers used to create the pQR plasmids containing with the full metabolic operon. The sequences include that which was used to generate the nested fragment templates which were then subsequently used in all other modification steps. Also included are the primers used to generate the individual his-tagged proteins for purification as well as the primers used to add in a second copy of *ChnD* T7 promoter, operator and terminator.

## A.1. Isolating the nested fragments and creating operon A and B

Table 8.1 contains the primers used to generate the nested fragments which were subsequently used as templates for all further modifications such as changing RBS, adding homologous overlaps and restrictions sites. The gene cluster sequence of 16.141 kb was accessed from NCBI available at:

https://www.ncbi.nlm.nih.gov/nuccore/AB006902.2?report=fasta&to=16160

| | |
|---|---|
| Nested *ChnA* for | 5' CTTCTTCTCAGGCAGATTCGCAG |
| Nested *ChnA* rev | 5' AGACGAGTCATCTACCGGATCTG |
| Nested ChnB for | 5' AAG CCT TGT GAT TGC ATT CCT GC |
| Nested ChnB rev | 5' TTA GGC ATT GGC AGG TTG CTT G |
| Nested ChnC for | 5' CGTGTGGGCACTGATTTATGCT |
| Nested ChnC rev | 5' GTATCAAACCAGTGATGCAAAGCC |
| Nested ChnD for | 5' CACTCAGGTCTTTGCGCAATTTC |
| Nested ChnD rev | 5' GATTCATGCACTGTTACTGCGTG |

**Table 8.1 A table of the primers used to generate the nested fragment from the metabolic cluster from *Acinetobacter calcoaceticus* NCIMB 9871**

The above primers created the following gene fragments when using a genomic sample from *Acinetobacter calcoaceticus* NCIMB 9871. The gene is in bold and underlined and the flanking nested portion is not. *ChnA, ChnC* and *ChnC* were in the reverse direction and so the above ''for'' primer would then become the reverse primer of the gene.

***ChnA* nested fragment**

AGA CGA GTC ATC TAC CGG ATC TGG TTT CTT GAA TCT TAA AAT ATG AAA CAA GCT TTG ATG CTT AAA ATT TGA AAT GGA AAA AAT T **ATG TCA AAT AAA TTC AAC AAT AAA GTC GCT TTA ATT ACT GGC GCT GGT TCA GGT ATT GGT AAA AGC ACC GCA CTG CTT TTG GCT CAA CAG GGT GTA AGT GTA GTG GTT TCA GAT ATT AAC CTG GAA GCA GCA CAG AAA GTT GTG GAC GAA ATT GTC GCT TTA GGC GGG AAA GCG GCT GCG AAT AAG GCC AAT ACT GCT GAG CCT GAA GAC ATG AAA GCT GCA GTC GAG TTT GCG GTC AGC ACT TTT GGT GCA CTG CAT TTG GCC TTC AAT AAT GCG GGA ATT CTG GGT GAA GTT AAC TCC ACC GAA GAA TTG AGC ATT GAA GGA TGG CGT CGT GTG ATT GAT GTG AAC TTG AAT GCG GTT TTC TAC AGC ATG CAT TAT GAA GTT CCT GCA ATC TTG GCC GCA GGG GGC GGA GCG ATT GTC AAT ACC GCT TCT ATT GCA GGC TTG ATC GGG ATT CAA AAT ATT TCA GGC TAT GTC GCT GCA AAA CAT GGC GTA ACG GGT CTA ACG AAA GCG GCG GCA TTG GAA TAT GCA GAT AAA GGG ATT CGC ATT AAT TCA GTA CAT CCT GGC TAT ATC AAA ACG CCT TTG ATT GCA GAA TTT GAA GAA GCA GAA ATG GTA AAA CTA CAT CCG ATT GGT CGT TTG GGA CAG CCG GAA GAA GTT GCT CAG GTT GTT GCC TTC CTA CTT TCT GAT GAT GCT TCA TTT GTG ACC GGT AGT CAG TAT GTG GTC GAT GGT GCA TAT ACC TCG AAA TAA** AAA ATG AGT AGA AGA TAC CCC CAT ATA AAG AGG AAA GAG GGG TGA TAC CCT CTT TTC TTT TAT CTA TTG TTG TAA GGG AAA ATT TTT CTT TGC TGC GAA TCT GCC TGA GAA GAA G

***ChnB* nested fragment**

AAG CCT TGT GAT TGC ATT CCT GCG ATT CTT TAT TCA ATG AAT AAG CAA TGC TAT TAA TCA GCA ATG AAT AAC CAG CAC TGC AGA TTT TGA ATA AAT TCA CAT GTC GTA ATG GAG ATT ATC **ATG TCA CAA AAA ATG GAT TTT GAT GCT ATC GTG ATT GGT GGT GGT TTT GGC GGA CTT TAT GCA GTC AAA AAA TTA AGA GAC GAG CTC GAA CTT AAG GTT CAG GCT TTT GAT AAA GCC ACG GAT GTC GCA GGT ACT TGG TAC TGG AAC CGT TAC CCA GGT GCA TTG ACG GAT ACA GAA ACC CAC CTC TAC TGC TAT TCT TGG GAT AAA GAA TTA CTA CAA TCG CTA GAA ATC AAG AAA AAA TAT GTG CAA GGC CCT GAT GTA CGC AAG TAT TTA CAG CAA GTG GCT GAA AAG CAT GAT TTA AAG AAG AGC TAT CAA TTC AAT ACC GCG GTT CAA TCG GCT CAT TAC AAC GAA GCA GAT GCC TTG TGG GAA GTC ACC ACT GAA TAT GGT GAT AAG TAC ACG GCG CGT TTC CTC ATC ACT GCT TTA GGC TTA TTG TCT GCG CCT AAC TTG CCA AAC ATC AAA GGC ATT AAT CAG TTT AAA GGT GAG CTG CAT CAT ACC AGC CGC TGG CCA GAT**

**GAC GTA AGT TTT GAA GGT AAA CGT GTC GGC GTG ATT GGT ACG GGT TCC ACC GGT GTT CAG GTT ATT ACG GCT GTG GCA CCT CTG GCT AAA CAC CTC ACT GTC TTC CAG CGT TCT GCA CAA TAC AGC GTT CCA ATT GGC AAT GAT CCA CTG TCT GAA GAA GAT GTT AAA AAG ATC AAA GAC AAT TAT GAC AAA ATT TGG GAT GGT GTA TGG AAT TCA GCC CTT GCC TTT GGC CTG AAT GAA AGC ACA GTG CCA GCA ATG AGC GTA TCA GCT GAA GAA CGC AAG GCA GTT TTT GAA AAG GCA TGG CAA ACA GGT GGC GGT TTC CGT TTC ATG TTT GAA ACT TTC GGT GAT ATT GCC ACC AAT ATG GAA GCC AAT ATC GAA GCG CAA AAT TTC ATT AAG GGT AAA ATT GCT GAA ATC GTC AAA GAT CCA GCC ATT GCA CAG AAG CTT ATG CCA CAG GAT TTG TAT GCA AAA CGT CCG TTG TGT GAC AGT GGT TAC TAC AAC ACC TTT AAC CGT GAC AAT GTC CGT TTA GAA GAT GTG AAA GCC AAT CCG ATT GTT GAA ATT ACC GAA AAC GGT GTG AAA CTC GAA AAT GGC GAT TTC GTT GAA TTA GAC ATG CTG ATA TGT GCC ACA GGT TTT GAT GCC GTC GAT GGC AAC TAT GTG CGC ATG GAC ATT CAA GGT AAA AAC GGC TTG GCC ATG AAA GAC TAC TGG AAA GAA GGT CCG TCG AGC TAT ATG GGT GTC ACC GTA AAT AAC TAT CCA AAC ATG TTC ATG GTG CTT GGA CCG AAT GGC CCG TTT ACC AAC CTG CCG CCA TCA ATT GAA TCA CAG GTG GAA TGG ATC AGT GAT ACC ATT CAA TAC ACG GTT GAA AAC AAT GTT GAA TCC ATT GAA GCG ACA AAA GAA GCG GAA GAA CAA TGG ACT CAA ACT TGC GCC AAT ATT GCG GAA ATG ACC TTA TTC CCT AAA GCG CAA TCC TGG ATT TTT GGT GCG AAT ATC CCG GGC AAG AAA AAC ACG GTT TAC TTC TAT CTC GGT GGT TTA AAA GAA TAT CGC AGT GCG CTA GCC AAC TGC AAA AAC CAT GCC TAT GAA GGT TTT GAT ATT CAA TTA CAA CGT TCA GAT ATC AAG CAA CCT GCC AAT GCC TAA**

***ChnC* nested fragment**

GTA TCA AAC CAG TGA TGC AAA GCC ACT CTT TAG ATT TGG ATC TTT GAT GGC ACT CAA GTT TTT ATA TCT TCA TGG AAT AAA ACT AGG ATT TAT TCC TTA AAT ATC AAA TCA AGG ATT TTA AA **A TGA ATA GCA CAC AAA GCA ATA CTC AAT TTC TTT TCG ATT TAT ATG CGA ACT GGT CAA GAC GGA TGC AGG AAA ATC CGA ATA TGA CCA TTG AAG ACT TTC GCA GTA TGT TTG ATG AAT GGC ATC AAC CTA CAT TGG AAC CGG AAG AAG TGT CTT ATA AAT TCG ATG TTG TGG CAG GTG TAG AAG GTC TTT GGA TTT ATC CGA AAG ATG CTG ACT TAT CCA AAG TCA TCA TTT ATA CCC ATG GCG GTG GAT TTG CGG TCG GTT CTT CGG CCA GTC ACC GTA AGC TGG TGG GGC ATT TGG CCA AGT ATT TAG GGG TAT CCG CAT TTG TGG TTG ATT ACC GAC GTT CAC CAG AAC ATG TCT TCC CGG CAC AAA TTC AGG ACG TGA CAG CAG TAT ATA AAG AAC TAC TCC AGC GTG GCT TTA CTG CAA AAA ATA TGC TGA CCG CAG GGG ATT CTG CGG GGG GGA ATC TGG CGA TAT CAA CCG TAC TCA ATC TAC GAA ATG AAG GGA TTG AGT**

**TGC CAG GAG CAG TGA TTG CAT TCT CTC CTT GGC TGG ATA TGG AGC ACA AAG GTG AAA CCC TGA TCA GCA ACG ATG CCA CTG ATG CCT TGA TTA CAG TGG ATC TGC TTA AAG GCA TGT CAC AAA TGT TCT TGG GTG AAC ATG GTG ATC CGG CAA ATC CAT TGG CGA ATC CGT TAA AAG CCA ATT ATC AGG TTT TCC CAC GTT TGT ATA TCA ATG CCG GAT CAG TTG AAT CAC TTG TAG ACA ATG CAA CAC GTC TTG CTG ATA TTG CAA AAA AAG AGG GTG TTG ATG TGA CTT TAT CTG TGG TGG ACA ACA TGC AGC ACG TTT TTC CTT TCC TAG CTG GGC GTG CAA GTG AAG CTG ATC AAG AAT TAG CGA AAA TTG CGC AGT GGT TTA AAG CAT AA**A TCA GTG CCC ACA CG

***ChnD* nested fragment**

GAT TC**A TGC ACT GTT ACT GCG TGA CGC ATC ATG GAC AAC CAC TCG AAG ACG TTG AGA AAG AAA TTC CGC AAC CGA AAG GTA CTG AAG TTT TAC TCC ATG TAA AAG CCG CAG GTC TAT GCC ATA CGG ATT TAC ACT TAT GGG AAG GTT ATT ATG ATC TAG GTG GGG GCA AGC GTT TAT CCC TTG CAG ATC GTG GGC TGA AGC CAC CCT TAA CCT TAA GTC ATG AAA TTA CAG GTC AGG TGG TTG CTG TCG GTC CAG ATG CGG AAT CAG TCA AGG TCG GCA TGG TCA GCT TGG TTC ATC CAT GGA TTG GTT GCG GTG AAT GCA ACT ACT GTA AAC GTG GCG AAG AAA ACC TGT GTG CCA AAC CGC AAC AGT TAG GCA TCG CCA AGC CGG GTG GTT TTG CCG AAT ATA TCA TCG TGC CGC ATC CAC GAT ATC TGG TGG ATA TTG CAG GTC TGG ATC TGG CTG AAG CTG CAC CTT TGG CAT GTG CAG GCG TGA CAA CAT ACA GTG CAC TGA AAA AAT TCG GTG ATT TGA TTC AAA GCG AGC CGG TGG TGA TCA TTG GTG CCG GTG GTT TAG GGC TGA TGG CAC TCG AGT TGC TCA AAG CTA TGC AAG CCA AAG GCG CAA TCG TAG TTG ATA TTG ATG ACA GCA AAC TGG AAG CAG CAC GTG CTG CCG GTG CAT TAT CGG TCA TCA ATA GCC GAA GTG AGG ATG CTG CTC AAC AGC TGA TTC AGG CAA CTG ACG GTG GTG CAC GTC TGA TCC TTG ATC TGG TTG GCA GTA ATC CAA CAT TGA GCC TTG CCT TGG CGA GTG CTG CAC GTG GTG GGC ATA TTG TGA TCT GCG GAT TGA TGG GGG GAG AAA TTA AGC TTT CCA TTC CGG TGA TTC CAA TGA GAC CAC TCA CAA TCC AGG GCA GTT ATG TAG GGA CGG TAG AGG AAT TAA GAG AGC TGG TGG AGC TGG TGA AAG AAA CCC ACA TGT CAG CCA TTC CCG TGA AAA AAC TGC CAA TTT CGC AGA TCA ATT CCG CAT TTG GAG ACT TGA AAG ATG GCA ACG TCA TCG GGC GTA TTG TGC TTA TGC ACG AAA ACT GA** T CAC CCA ACG TTT TTC CAT ATT TAA TAC CCA AAT CAA AAT CAA GAA ATT GCG CAA AGA CCT GAG TG

The following Table (Table 8.2) contains the primers used to create the gene fragments for the CPEC and gibson assembly reactions to create operon A (native RBS). pET29a was inverse PCR amplified to linearise it and create overlaps to *ChnA* and *ChnD.*

| | |
|---|---|
| OPA *ChnA* for Vector overlap (with *NdeI* site) | AAGAAGGAGATATA*CATATG***GGA**AAAAATT**ATGTCAAATAAATTCAACAATAAAGTC** |
| OPA *ChnA* rev B ov | CCATTTTTTGTGACATGATAA**TCTCCTTATTTCGAGGTATATGCACCAT** |
| OPA ChnB for A ov | ATGGTGCATATACCTCGAAATAA**GGAGA**TTATC**ATGTCACAAA** |
| OPA ChnB rev C ov | GTGTGCTATTCATTTTAAAA**TCCTTTTAGGCATTGGCAGGTTGC** |
| OPA ChnC for B ov | GCAACCTGCCAATGCCTAA**AAGGA**TTTTAAA**ATGAATAGCACAC** |
| OPA ChnC rev D ov | CACGCAGTAACAGTGCATGAATC**TCTCCTTTTATGCTTTAAACCACTGCGCAATT** |
| OPA ChnD for C ov | AATTGCGCAGTGGTTTAAAGCATAA**AAGGAGA**GATTC**ATGCACTGTTACTGCGTG** |
| OPA ChnD rev Vector overlap (with *NotI* site) | TGGTGCTCGAGT**GCGGCCGC**TCAGTTTTCGTGCATAAGCACAATACG |
| OPA pET29a inverse PCR rev *ChnA* overlap | GAATTTATTTGACATAATTTTT**TCC***CATATG***TATATCTCCTTCTTAAAG** |
| OPA pET29a inverse PCR for ChnD overlap | GTATTGTGCTTATGCACGAAAACTGA*GCGGCCGC***ACTCGAGCACCACCA** |

**Table 8.2 Primers to create the fragments for operon A; native RBS on the *Acinetobacter* genes. In italics and underlined are the restriction sites (*NdeI* on *ChnA* and *NotI* on *ChnD*). In bold and underlined is the native RBS used in operon A and in bold only is the gene. The 5' of the forward primer contains a stretch of nucleotides homologous to the 5' flanking end of the gene e.g. in the case of *ChnA* it would be the pET29a vector. These are immediately before or after the restriction site or RBS. The reverse primer has the homologous region following the end of the gene.**

The following table (Table 8.3) are the primers used to create the artificial RBS and overlaps for the CPEC assembly into operon B. The template was each respective nested fragment. The Tm was calculated based only on the gene binding portion and not the overlaps. pET29a was inverse PCR amplified to linearise it and create overlaps to *ChnA* and *ChnD.* OPA pET29a inverse PCR for ChnD overlap was used in Operon B on pET29a since this portion of the plasmid is identical in both versions of the operon.

| | |
|---|---|
| OPB *ChnA* for Vector overlap (with *NdeI* site) | TTAAGAAGGAGATATA*CATATG***AGGAG**AAAATT**ATGTCAAATAAATTCAACAATAAAGTC** |
| OPB *ChnA* rev B ov | CCATTTTTTGTGACATGATAAA**CTCCTTTATTTCGAGGTATATGCACCAT** |
| OPB ChnB for A ov | ATGGTGCATATACCTCGAAATAA**AGGAG**TTTATC**ATGTCACAAAAAATGG** |
| OPB ChnB rev C ov | GCTTTGTGTGCTATTCATTTTAAA**CTCCTTTAGGCATTGGCAGGTTGCTTG** |
| OPB ChnC for B ov | CAAGCAACCTGCCAATGCCTAA**AGGAG**TTTAAA**ATGAATAGCACACAAAGC** |
| OPB ChnC rev D ov | CACGCAGTAACAGTGCATGAATCA**CTCCTTTATGCTTTAAACCACTGCGCAA** |
| OPB ChnD for C ov | TTGCGCAGTGGTTTAAAGCATAA**AGGAG**TGATTC**ATGCACTGTTACTGCGTG** |
| OPB ChnD rev Vector overlap (with *NotI* site) | GTGGTGGTGGTGCTCGAGT*GCGGCCGC***TCAGTTTTCGTGCATAAGCAC** |
| OPB pET29a inverse PCR rev *ChnA* overlap | GAATTTATTTGACATAATTTT**CTCCT***CATATG***TATATCTCCTTCTTAAAG** |
| OPA pET29a inverse PCR for ChnD overlap | GTATTGTGCTTATGCACGAAAACTGA*GCGGCCGC***ACTCGAGCACCACCA** |

**Table 8.3 Primers to create the fragments for operon B; artificial RBS on the *Acinetobacter* genes. In italics and underlined are the restriction sites (*NdeI* on *ChnA* and *NotI* on *ChnD*). In bold and underlined is the artificial RBS used in operon A and in bold only is the gene. The 5' of the forward primer contains a stretch of nucleotides homologous to the 5' flanking end of the gene e.g. in the case of *ChnA* it would be the pET29a vector. These are immediately before or after the restriction site or RBS. The reverse primer has the homologous region following the end of the gene.**

## A.2. Cloning the ω-TAm into both operon versions

As stated in the results section, the PP_5182 TAm from *Pseudomonas putida* KT2440 and Dgeo_2743 from *Deinococcus geothermalis* were both cloned using CPEC. The primers for doing so are found in Table 8.4. The DNA sequences for both genes are listed below and obtained from the KEGG database. The OpA/B primers were used to inverse PCR both operon versions and removing the *NotI* site. Both TAm's used the AGGAG RBS and a *BamHI* site was introduced as the 6 nucleotide spacer between the RBS and the TAm. the linearised vector was then used as a template with subsequent primers to add on homologous overlaps for the TAm. The TAm was then amplified with it's respective overlaps, RBS and *BamHI* site.

**pQR810: *Pseudomonas putida* KT2440** gene construct **PP_5182**

**DNA sequence:**

atgagcgtcaacaacccgcaaacccgtgaatggcaaaccctgagcggggagcatcacctc
gcacctttcagtgactacaagcagctgaaggagaaggggccgcgcatcatcaccaaggcc
cagggtgtgcatttgtgggatagcgaggggcacaagatcctcgacggcatggccggtcta
tggtgcgtggcggtcggctacggacgtgaagagctggtgcaggcggcggaaaaacagatg
cgcgagctgccgtactacaacctgttcttccagaccgctcacccgcctgcgctcgagctg
gccaaggcgatcaccgacgtggcgccgaaaggtatgacccatgtgttcttcaccggctcc
ggctccgaaggcaacgacactgtgctgcgcatggtgcgtcactactgggcgctgaagggc
aaaccgcacaagcagaccatcatcggccgcatcaacggttaccacggctccaccttcgcc
ggtgcatgcctgggcggtatgagcggcatgcacgagcagggtggcctgccgatcccgggc
atcgtgcacatccctcagccgtactggttcggcgagggaggcgacatgacccctgacgaa
ttcggtgtctgggccgccgagcagttggagaagaagatcctcgaagtcggcgaagacaac
gtcgcggccttcatcgccgagccgatccagggcgctggtggcgtgatcatcccgccggaa
acctactggccgaaggtgaaggagatcctcgccaggtacgacatcctgttcgtcgccgac
gaggtgatctgcggcttcggccgtaccggcgagtggttcggctcggactactacgacctc
aagcccgacctgatgaccatcgcgaaaggcctgacctccggttacatccccatgggcggt
gtgatcgtgcgtgacaccgtggccaaggtgatcagcgaaggcggcgacttcaaccacggt
ttcacctactccggccacccggtggcggccgcggtgggcctggaaaacctgcgcattctg
cgtgacgagaaaattgtcgagaaggcgcgcacggaagcggcaccgtatttgcaaaagcgt
ttgcgcgagctgcaagaccatccactggtgggtgaagtgcgcggcctgggcatgctggga

gcgatcgagctggtcaaggacaaggcaacccgcagccgttacgagggcaagggcgttggc
atgatctgtcgcaccttctgcttcgagaacggcctgatcatgcgtgcggtgggtgacacc
atgatcatcgcgccgccgctggtaatcagccatgcggagatcgacgaactggtggaaaag
gcgcgcaagtgcctggacctgacccttgaggcgattcaataa

**pQR983:** ***Deinococcus geothermalis*** gene construct **Dgeo_2743**

**DNA sequence:**

atgagcaagaatatcgattggcagaacctgggcttcagctacatcaagaccgacctgcgc
taccgctcgcactggaaggatggcgagtgggacgcaggggtgctcaccgaggacaactgc
gttcacatcagcgagggctcgaccgctctgcactacgggcagcagtgctttgagggtctc
aaggcataccgccagcgcgatgggtcggtggtcctctttcgccccgaccagaacgcggcg
cggatgcaggcgagctgcgagcggctgctgatgccgaccatttcgaccgagcagttcatc
gctgcttgcgtgcaggtcgtgaaggccaacctggactatctgcctccctacggcacgggc
ggcgccctctacctgcggccctacatgatcggtgtgggcgacaacctgggcgtacgcacc
gccccggagttcctcttctccgtgttttgcgtaccggttggcgcctactttaagggcggc
ctgactcctgccaactttatcgtttcgccctacgaccgggccgcaccccacggcaccgga
gcggccaaggtgggcgggaactatgctgcgagcctcctccccggccacctcgccaaggaa
cgcggctttgccgacgcgatctacctcgatccacagacccataccaagatcgaggaggtg
ggggcggccaacttcttcgccatcacgcaggatgaccgcttcgtcacgccgcagtcgccc
tccatcctccccagcatcaccaagttctcgctgctgcagctggcaagggagcggctgggc
ctggaggtggaggaaggagacatctacatcgatcagctcgaccggtacaaggaggcggga
gcctgcggcacggcggcggtgatcacgcccatcggcggcatccagtacggtgaccacttc
cacgtcttttacagtgaaacggaggtcggcccggtgacccggcggctgtatgatgagttg
gttggcatccagttcggggacatcgcggccccagagggctggctcgtccgagtggagtag

| | |
|---|---|
| OpA/B for | TCAGTTTTCGTGCATAAGCACAATACGC |
| OpA/B rev | GCTGCTAACAAAGCCCGAAAGG |
| CPEC 810 rev | GTTGTTGACGCTCAT*GGATCC***CTCCT**TCAGTTTTCGTGCATAAGCACAATACGC** |
| CPEC 810 for | GAGGCGATTCAATAA*GGATCC***GCTGCTAACAAAGCCCGAAAGG** |
| CPEC 983 rev | GATATTCTTGCTCAT*GGATCC***CTCCT**TCAGTTTTCGTGCATAAGCACAATACGC** |
| CPEC 983 for | GAGTGGAGTAG*GGATCC***GCTGCTAACAAAGCCCGAAAGG** |

**Table 8.4 Primers used to inverse PCR both operon A and B plasmids, removing the *NotI* site and adding in overlaps for each TAm. In italics and underlined is the restriction site *BamHI*. In bold and underlined is the artificial RBS used for both TAm's. The 5' of the forward primer contains a stretch of nucleotides homologous to the 5' flanking end of the gene that end is going to be placed next too.**

Table 8.5 lists the primers used on plasmid minipreps of each TAm to generate the fragments required for the CPEC reaction.

| | |
|---|---|
| pQR810 ChnD o/v for | GTGCTTATGCACGAAAACTGA**AGGAG***GGATCC***ATGAGCGTCAACAACCCGCAAAC** |
| pQR810 vec o/v rev | CCTTTCGGGCTTTGTTAGCAGC*GGATCC* **TT ATT GAA TCG CCT CAA G** |
| pQR983 ChnD o/v for | GTGCTTATGCACGAAAACTGA**AGGAG***GGATCC***ATGAGCAAGAATATCGATTGG** |
| pQR983 vec o/v rev | CCTTTCGGGCTTTGTTAGCAGC*GGATCC* **CTA CTC CAC TCG GAC GAG** |

**Table 8.5 PCR primers used on plasmid minipreps from pQR810 and pQR983 to add on a *ChnD* and vector overlap, RBS and *BamHI* site to each TAm.**

After unsuccessfully attempting to clone pQR1014 by CPEC, it was cloned by standard restriction cloning using *AvrII* and BbvCI sites into operon A only. The RBS was added directed after *ChnD* which was immediately followed by the *AvrII* site. The BbvCI site was added to the 3' end of the MLUT_00920 (sometimes identified as MLUT_RS12025).

The DNA sequence for MLUT_00920 (aka MLUT_RS12025) is as follows:

ATGTCCGAGCATCAGCAGCACGTCGGCCGCCGGGACCCCCGGATCAGCGACGAGGCCATCGAGCG
GGGCCGCCGCGCCCACGAGCTCGACCGCGCGCACGTGTTCCACTCGTGGAGCGCGCAGCGCGATC
TGAATCCGATGACCATCCTCGACGCCGAGGGGTCGTGGGTGTGGGACGGTGAGGGTCGGCCCATG
GTCGACTTCAGCGGGCAGCTCGTCTTCACGAACGTCGGCCACCGCCACCCGAAGATCGTGGCCGCG
ATCCGCGAACAGGCCGAGACGCTCGCCACCGTGGCCCCGCAGCACGTCAACGACGCCCGGTCCGA
GGCAGCCCGGCTGATCACCGAGCGGCTGCCGGAGTCCATCAACCACGTGTTCTTCACGAACGCGG
GCGCGGAGGCCAACGAGCACGCGGTGCGGATGGCACGCCTGCACACGGGTCGGCACAAGGTGCT
CTCCGCGTACCTCTCGTACCACGGCGCCACGCGCCTGACCGCGAACCTGACCGGCTCCCTGCGACG
CGTGGGCTCGGACTCGGCGTCCGACGGCGTCGTGCACTTCCAGCCCGCGTACACGTACCGCTCCGC
GTTCGGCTCGGAGAGCGATGAGCAGGAGGCCGAGCGCGCGCTGGCCCACCTGCGCGACGTGATC
GAGCTCGAGGGTCCCTCCACCGTGGCCGCCGTGGTCCTCGAGGCCGTCCCCGGCACCGCGGGCAT
CTTCCTGCCGCCGCCGGGCTACATGGCCGGCGTCGGGGAGCTGTGCCGCGAGCACGGGATCCTGC
TGATCATCGACGAGGTCATGGTCGGCTTCGGCCGCGTGGGCGAGTGGTTCGGCCACCAGCTCACG
GGCGTCACCCCGGACATCGTCACGTTCGCCAAGGGGGTCAACTCCGGGTACGTGCCGTTGGGCGG
CGTCGCCATGAGCGACGACGTCTACGAGAGCTTCGCGACCACGCCCTACCCCGGCGGCCTGACCTA
CTCGGGCCACCCCCTCGCGTGCGCGGCCGCCGTCGCCGCGATCACCGCGATGGAGGAGGAGGGCA
TGGTGGCCCATGCGAAGCGCATCGGCGAGGAGGTCATCGGGCCGCGCCTGGCCGAGATCGCCGT
GGCCCATCCGAGCGTGGGCGACGTGCGCGGCGCCGGCGCGTTCTGGGCCGTCGAGCTGGTGAAG
GACCGGCAGACCCGGGAGCGGTTGGCCCCGCTCGGCCAGGTGGCGCCCGTGATGGGGCGGATGA
TGGCCGAGGCCAAGGACCGCGGGCTGCTGCTGTTCATGGCCGAGAACCGGTTCCACCTGTGCCCG
CCGCTGAACATCTCGGACGAGGACCTGCGCTTCGGCCTCGACGTGCTCGACCGGGTCCTCACCCTC
GCCGACGAGGAGGTCGCGGCCAGCTGA

Table 8.6 contains the primers used in the cloning of MLUT_00920 into operon A only. The RBA was added onto the linearised operon A vector followed by the *AvrII* site.

| | |
|---|---|
| OPA rev *AvrII* | GGTTAT*CCTAGG***CTCCTTCAGTTTTCGTGCATAAGCACAATA** |
| OPA for BbvCI | ATCTAG*CCTCAGC***GCTGCTAACAAAGCCCGAAAGG** |
| 1014 for (*AvrII*) | AACTAT*CCTAGG***ATGTCCGAGCATCAGCAGCAC** |
| 1014 rev (BbvCI) | CGATTA*GCTGAGG***TCAGCTGGCCGCGACCTC** |

**Table 8.6 Primers used in the cloning of MLUT_00920 into operon A**

## A.3. Individually cloning the *Acinetobacter calcoaceticus* NCIMB 9871 genes individually

Each gene taken from *Acinetobacter* was cloned individually into a pET29a vector. Each genes stop codon was removed and the RBS from the vector as well as the his-tag was used for each gene. The *NdeI* site was positioned at the 5' end and the *NotI* site at the 3' end of the gene. The pET29a vector was linearised by double digestion with both enzymes followed by gel purification.

| | |
|---|---|
| *ChnA* for idvc (*NdeI*) | ATCGTA*CATATG***ATGTCAAATAAATTCAACAATAAAG** |
| *ChnA* rev idvc (*NotI*) | TAACTA*GCGGCCGC***TTTCGAGGTATATGCACC** |
| ChnB for idvc (*NdeI*) | ATCGTA*CATATG***ATGTCACAAAAAATGGATTTTGATGC** |
| ChnB rev idvc (*NotI*) | TAACTA*GCGGCCGC***GGCATTGGCAGGTTGCTT** |
| ChnC for idvc (*NdeI*) | ATCGTA*CATATG***ATGAATAGCACACAAAGCAATAC** |
| ChnC rev idvc (*NotI*) | TAACTA*GCGGCCGC* **TGC TTT AAA CCA CTG C** |
| ChnD for idvc (*NdeI*) | ATCGTA*CATATG***ATG CAC TGT TAC TGC GTG AC** |
| ChnD rev idvc (*NotI*) | TAACTA*GCGGCCGC* **GTT TTC GTG CAT AAG C** |

**Table 8.7 Primers used for individually cloning each enzyme into a pET29a vector for his-tag purification.**

## A.4. Assembling pQR1559: Cloning a T7-operator-*ChnD*-terminator sequence into pQR1553

The addition of a second copy of *ChnD* regulated by its own promoter, operator and terminator required several rounds of iteration. The fragment called *AvrII*-T7-operator-*ChnD*-terminator-*SalI* was inserted into a inverse PCR lineaerised pQR1553 plasmid at a site between the ROP and LacI sequences and the *AvrII* and *SalI* sites were added. The nested fragment of *ChnD* was used to generate this fragment. The two fragments were restriction cloned together.

| Step 1 | Nested fragment |
|---|---|
| AvrIT7opChnD for *AvrII* | CTAGTC*CCTAGG*TAATACGACTCACTATAGG*GGAATTGTGAGCGGATAACAATTCC****AAGG AGA*****GATTATGCACTGTTACTG** |
| ChnD rev *SalI* | TAGCTA*GTCGAC***TCAGTTTTCGTGCATAAGCACAATACGCCCGATGACGTT GCC ATC** |
| Step 2 | Using step 1 as a template |
| T7ChnDterm rev | CAA AAA ACC CCT CAA GAC CCG TTT AGA GGC CCC AAG GGG TTA TGC TAG **TCA GTT TTC GTG CAT AAG CAC AAT AC** |
| AvrIT7opChnD for | |
| Step 3 | Product of step 2 |
| ChnDfinalRev *SalI* | TAGCTA*GTCGAC*CAAAAAACCCCTCAAGACCCGTTTAGAG |
| ChnDfinalFor *AvrII* | CTAGTC*CCTAGG*TAATACGACTCACTATAGGGGAATTGTGAG |
| | |
| | Inverse PCR of pQR1553 |
| 53rev*AvrII* | ACGTTA*CCTAGG*GTCTGCGACCTGAGCAACAACATGAATG |
| 53for*SalI* | TCGATG*GTCGAC*GTATCGGTGATTCATTCTGCTAACCAGTAAG |

**Table 8.8 A list of the primers used to create pQR1559.**

# Appendix B

Table 8.9 is a list of the plasmid constructs and what each contain.

| | |
|---|---|
| pQR1543 | **pET29a encoding *ChnA-D* under native RBS sequences** |
| pQR1548 | **pET29a encoding *ChnA-D* under artificial RBS sequences** |
| pQR1553 | **pQR1543 + *PP5182*** |
| pQR1554 | **pQR1548 + *PP5182*** |
| pQR1555 | **pQR1543 + *DGEO_2743*** |
| pQR1556 | **pQR1548 + *DGEO_2743*** |
| pQR1557 | **pQR1543 + *MLUT_00920*** |
| pQR1559 | **pQR1553 + T7-operator-*ChnD*-terminator** |
| pQR426 | ***Pseudomonas aeruginosa* PAO2, PAO132** |
| pQR427 | ***Pseudomonas putida* KT2440, PP0596** |
| QR803 | ***Saccharopolyspora erythrea* DSM40517 Sery1824 = SACE3329** |
| pQR804 | ***Saccharopolyspora erythrea* DSM40517 Sery1599 = SACE4673** |
| pQR805 | ***Saccharopolyspora erythrea* DSM40517 Sery1902 = SACE5670** |
| pQR808 | ***Streptomyces coelicolor* A3(2) DSM40783 Sco5655** |
| pQR809 | ***Thermobifida fusca* DSM43792 Tfus0278** |
| pQR810 | ***Pseudomonas putida* KT2440 PP5182** |
| pQR811 | ***Pseudomonas putida* KT2440 PP2799** |
| pQR812 | ***Pseudomonas putida* KT2440 PP2588** |
| pQR813 | ***Pseudomonas aeruginosa* PAO2 PAO221** |
| pQR815 | ***Pseudomonas aeruginosa* PAO2 PA5313** |
| CV2025 | ***Chromobacterium violaceum* DSM30191** |
| pQR965 | ***Saccharopolyspora erythraea* SACE_0784** |
| pQR966 | ***Saccharopolyspora erythraea* SACE_1875** |
| pQR967 | ***Saccharopolyspora erythraea* SACE_2128** |
| pQR972 | ***Saccharopolyspora erythraea* SACE_3766** |
| pQR974 | ***Saccharopolyspora erythraea* SACE_6022** |
| pQR979 | ***Deinococcus geothermalis* Dgeo_1219** |
| pQR983 | ***Deinococcus geothermalis* Dgeo_2743** |
| pQR985 | ***Bacillus licheniformis* BLi00474** |
| pQR986 | ***Bacillus licheniformis* BLi00767** |
| pQR987 | ***Bacillus licheniformis* BLi01183** |
| pQR1003 | ***Vibrio fluvialis*** |
| pQR1010 | ***Klebsiella pneumoniae* KPN_01493** |
| pQR1011 | ***Klebsiella pneumoniae* KPN_03745** |
| pQR1014 | ***Micrococcus luteus* Mlut_00920** |
| pQR1015 | ***Micrococcus luteus* Mlut_10360** |

**Table 8.9 The plasmid construct I.D and what each is encoding in a pET29a vector**

# Appendix C

Figure 8.2 shows a typical chromatograph and spectra produced from lysed cell cultures. The peaks were compared to standards and quanitified using standard curves (Figure 8.3 - Figure 8.8). Retention times and peak separation was optimised by chanign column flow rate, run duration and split flow.

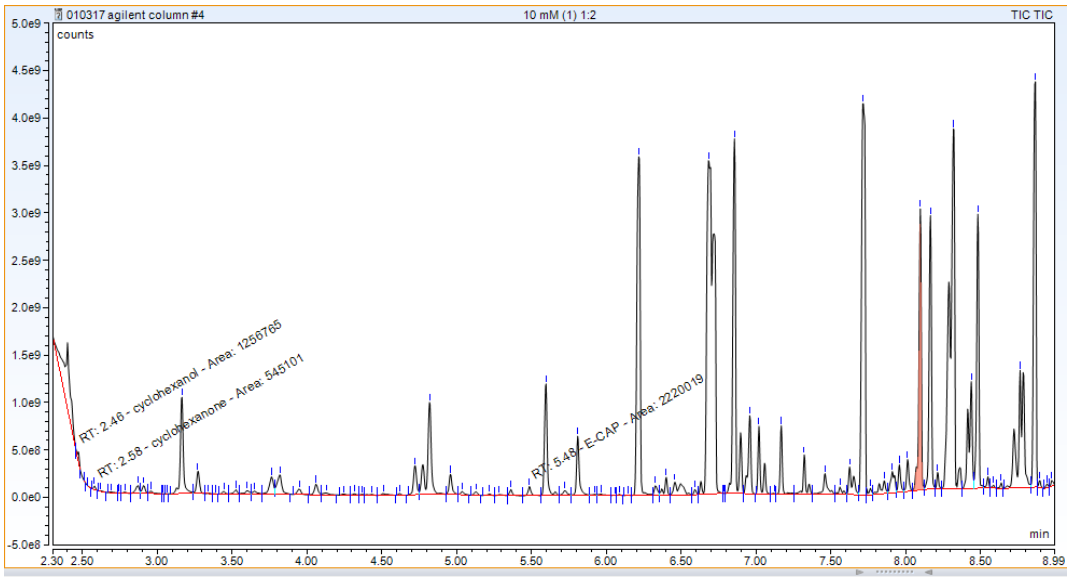



**Figure 8.2 Above is a typical chromatograph, diluted 1:2 in ethyl acetate of dried cell lysate from a 10 mM cyclohexanol culture. Below is the spectra of the peak at 8.1 min corresponding to 6-ACA**

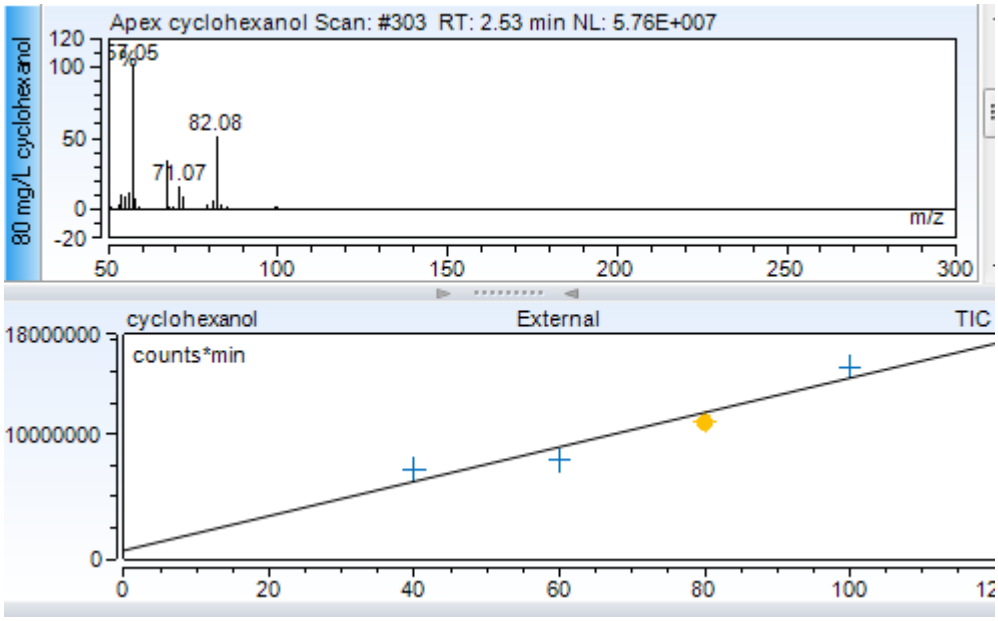



Figure 8.3 Cyclohexanol standard curve and spectra

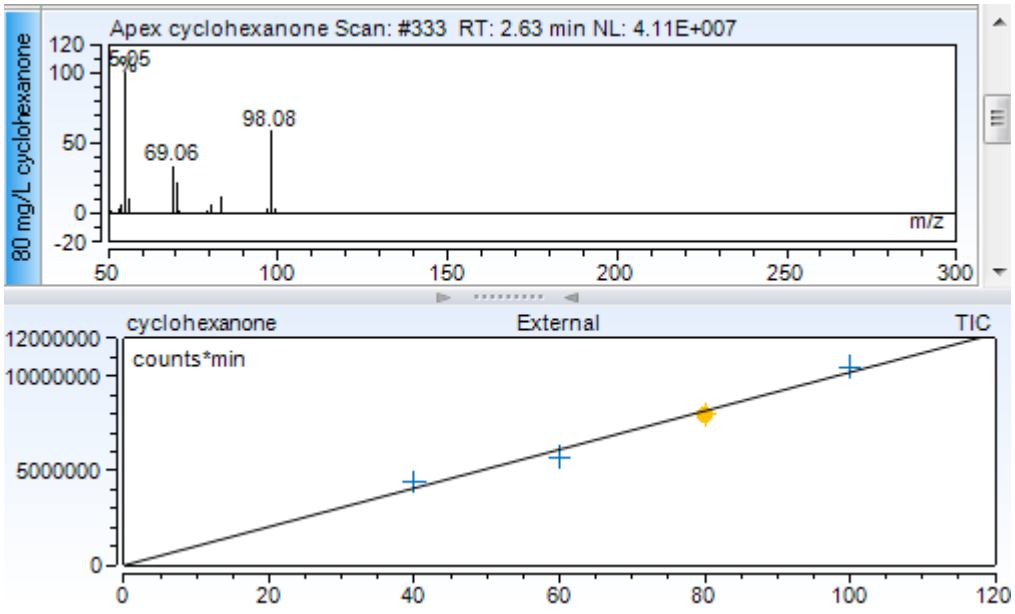



Figure 8.4 Cyclohexanone standard curve and spectra from GC-MS

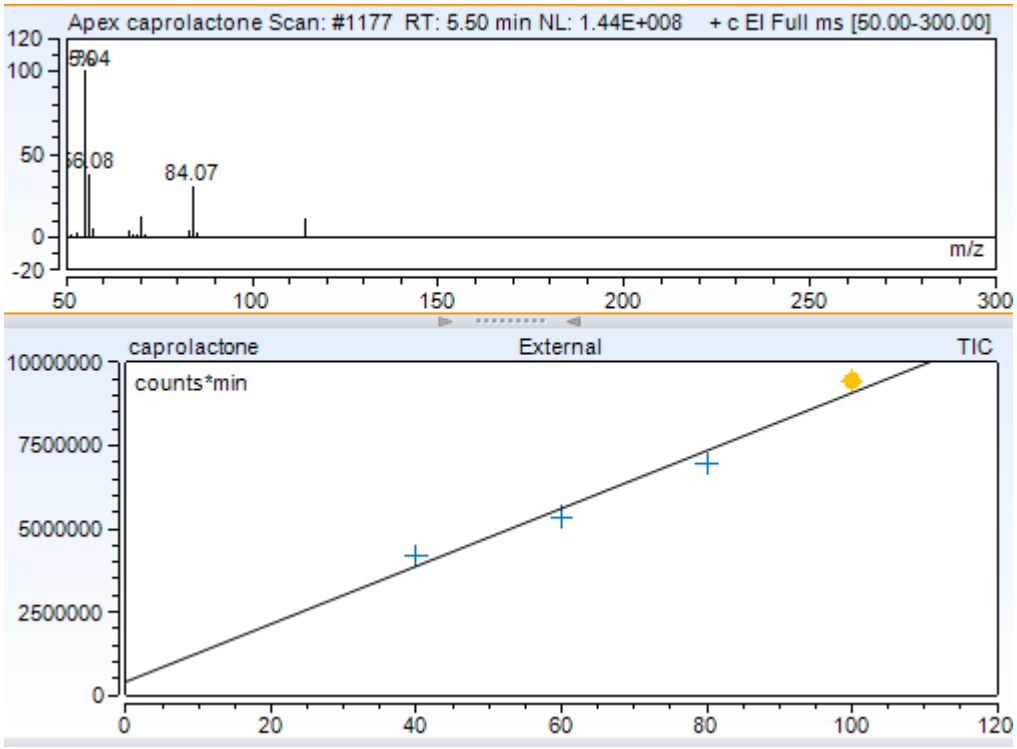



Figure 8.5 Caprolactone standard curve and spectra from GC-MS

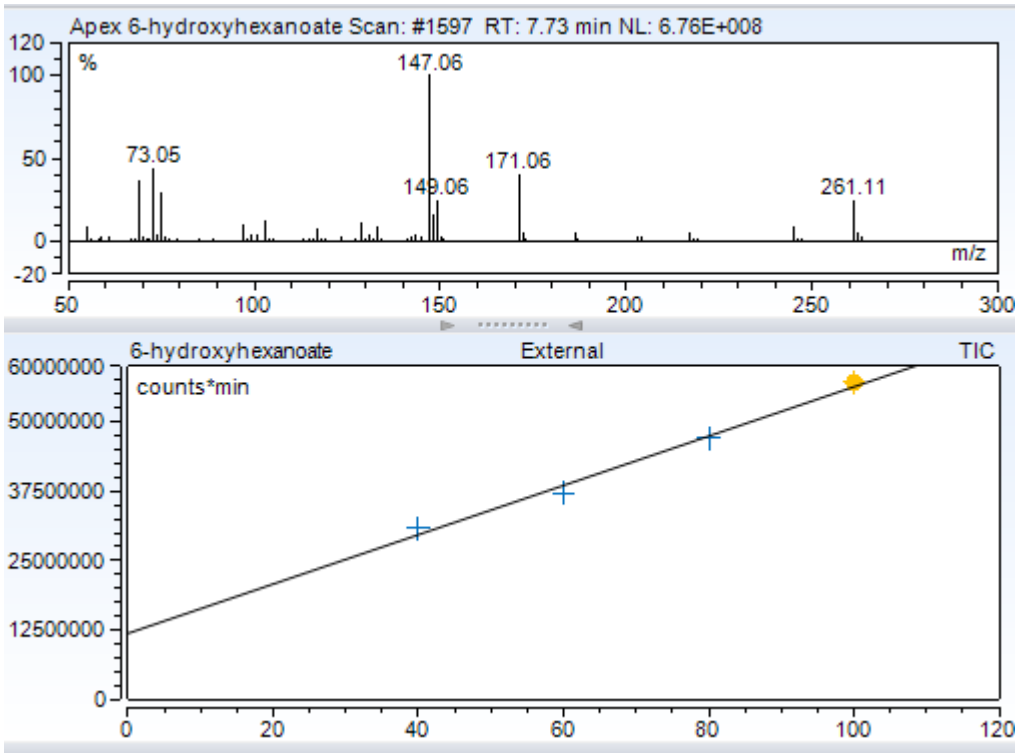



Figure 8.6. 2-TMS-6-hydroxyhexanoate standard curve and spectra from GC-MS

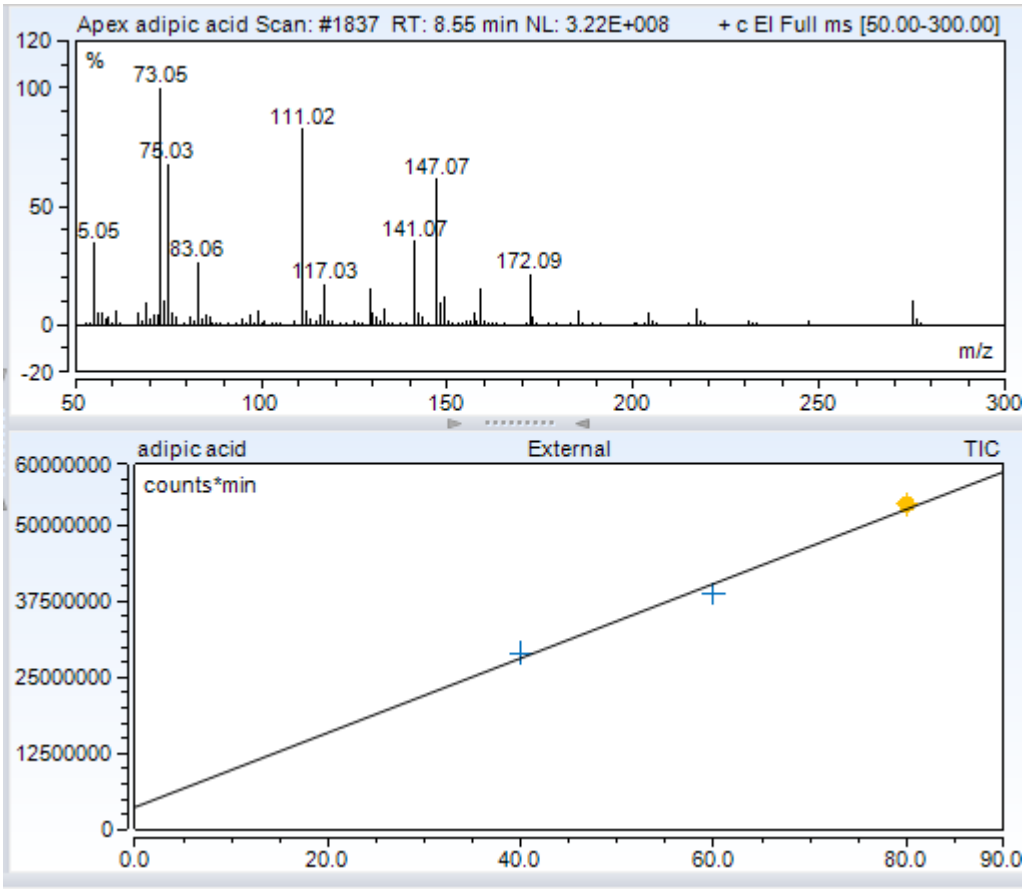
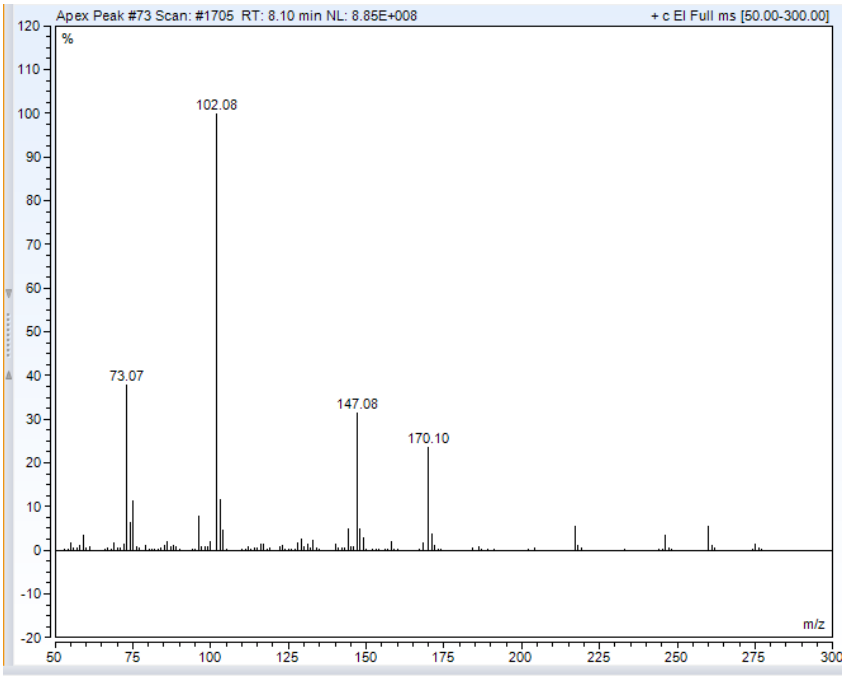



**Figure 8.7 Adipic acid standard curve and GC-MS spectra**



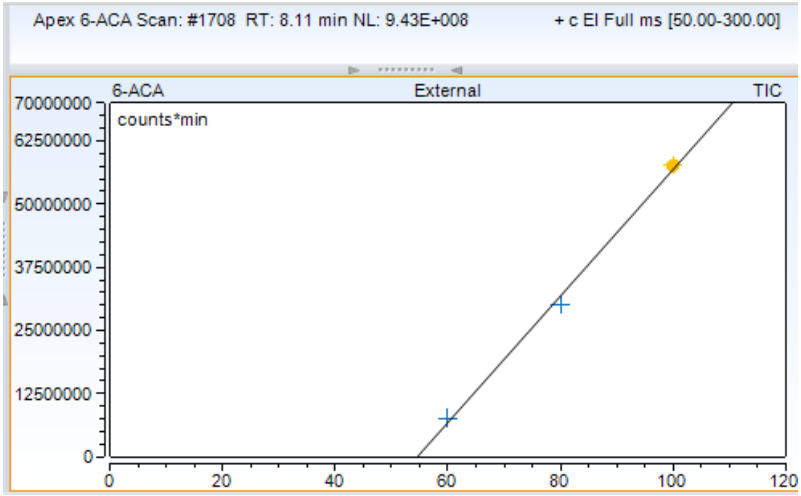




**Figure 8.8. 2-TMS-6Aminocaproic acid GC-MS spectra and standard curve**