Sustainable Biocatalysis: Exploring the CO_2 fixation activity of an alpha-keto acid decarboxylase

By

Parisa Hamzavinejad Moghaddam

School of Chemistry, Biodiscovery Institute

University of Nottingham, UK

Submitted for the degree for

Doctor of Philosophy

March 2021

Acknowledgment

Publication

- Biocatalytic synthesis of the amino acid L-homoalanine by exploiting the CO_2 fixation activity of an a-keto decarboxylase**.** P. Hamzavinejad moghaddam, D. Sagnelli, S. M. Howdle, N. R. Thomas to be submitted to *Green Chemistry*.

Abstract

Carbon dioxide (CO_2) is the major greenhouse gas formed from fossil fuels. The development of efficient and sustainable CO_2 fixation reactions for the conversion of this waste gas into useful chemicals is one of the most important challenges of current chemistry. The aim of this thesis is to show the application of pyruvate decarboxylase from yeast, *Saccharomyces cerevisiae* (*Sc*PDC) for the fixation of CO_2 on its own or in an artificial enzyme cascade to produce key chemical platforms. Firstly, *Sc*PDC was cloned, expressed and purified. The carboxylase activity of *Sc*PDC to convert aldehydes to their corresponding 2-ketoacids was investigated with the carbonate buffer providing a source of CO_2. Since the carboxylation reaction of *Sc*PDC in the carbonate buffer failed to generate any 2-ketoacids, namely pyruvate or 2-ketobutryic acid, branched chain decarboxylase from *Lactococcus lactis*, *Ll*KdcA was employed. *Ll*KdcA also failed to show any carboxylase activity in carbonate buffer. To overcome the unfavoured carboxylation reaction and shift the position of the equilibrium towards 2-ketoacid formation, a transaminase enzyme was coupled to the carboxylation reaction. In the cascade the transaminase enzyme should selectively aminate the product of the carboxylation reaction, and not the aldehyde which is the substrate for the carboxylation reaction. For this purpose, commercially available ω-transaminases were screened with different aldehydes and their corresponding ketoacids. Due to lack of selectivity observed with ω-transaminases, alanine transaminase from *Saccharomyces cerevisiae* (*Sc*ALT) was cloned, expressed and purified. A one-pot biocatalytic approach for the synthesis of L-alanine and L-homoalanine from CO_2 and their corresponding aldehydes was developed. Our finding showed that the carbonate buffer alone is not a sufficient source of CO_2 to achieve the reverse reaction of *Sc*PDC. In fact, addition of pressurised CO_2 (70psi, 0.48 MPa) to the reaction resulted in the production of L-alanine (0.040 mM) and L-homoalanine (0.262 mM). A further improvement in the yield of L-homoalanine (0.683 mM) was achieved by increasing the pH of the reaction (pH 9.5 at CO_2 pressure 70 psi, 0.48 MPa). The major limiting factor for the reaction going to completion (only 2.7% yield) was identified as the competing aldol condensation of the propanal that occurred under the preferred reaction conditions. Finally, *Sc*PDC was engineered by site-directed mutagenesis in order to expand its substrate specificity. Five active site variants of *Sc*PDC were generated based on the crystal structure of

the enzyme. One of these variants (T415A) showed enhanced specificity towards 2-ketobutryic acid in the decarboxylation direction.

Table of Contents

Table of Figures

Table of Schemes

List of Tables

Abbreviations

°C	Degree of Celsius
3-PG	3-phosphoglycerate
AADHs	Amino acid dehydrogenases
ADH	Alcohol dehydrogenase
ATP	Adenosine triphosphate
CO_2	Carbon dioxide
CODH	Carbon monoxide dehydrogenase
C. maltosa	*Candida maltosa*
DAD	Diode array detector
DNA	Deoxyribonucleic acid
*EC*IPDC	Pyruvate decarboxylase from *Enterobacter cloacae*
Expasy	Expert protein analysis system
FA	Furoic acid
FDCA	Furandicarboxylic acid
FDH	Formate dehydrogenase
FMOC	9-fluorenylmethyl chlorocarbonate
Gt	Giga tone
HPLC	High performance liquid chromatography
IMAC	Immobilised metal affinity chromatography
IPTG	Isopropyl-β-D-thiogalactopyranoside
kDa	Kilo dalton
Kivd	α-ketoisovalarate decarboxylase
KLD	kinase, ligase and DpnI
LB medium	Luria-Bertani medium
LDH	Lactate dehydrogenase
LeuDH	Leucine dehydrogenase
*Ll*KdcA	branched chain decarboxylase from *Lactococcus lactis*
MHz	Megahertz
MSP	Mitochondrial signal peptide
MWCO	Molecular weight cut off

NAD$^+$	Nicotinamide adenine dinucleotide (oxidised)
NADH	Nicotinamide adenine dinucleotide (reduced)
NMR	Nuclear magnetic resonance
OD$_{595}$	Optical density at the wavelength of 595nm
OPA	*O*-Phthaldialdehyde
Pa	Pascal
PCR	Polymerase chain reaction
PDC	Pyruvate decarboxylase
PLP	pyridoxal-5'-phoshpate
PMP	pyridoxamine-5'-phosphate
ppm	parts per million
prFMN	Prenylated flavin
RI	Refractive index
rpm	Revolutions per minute
RubisCO	Ribulose-1,5- carboxylate/oxygenase
RuBP	Ribulose-1,5-biphosphate
S	second
*Sc*ALT	Alanine transaminase from *Saccharomyces cerevisiae*
*Sc*PDC	Pyruvate decarboxylase from *Saccharomyces cerevisiae*
SDM	Site directed mutagenesis
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
*Su*PDC	Pyruvate decarboxylase from *Saccharomyces uvarum*
T$_a$	Annealing temperature
TAs	Transaminases
TCA	Tricarboxylic acid
ThDP	Thiamine diphosphate
U	Units
UV	Ultraviolet
WT	Wild type
YbdL	Methionine aminotransferase
*Zm*PDC	Pyruvate decarboxylase from *Zymomonas mobilis*

α	Alpha
β	beta
γ	Gamma
δ	Chemical shifts
ε	Extinction coefficient
μm	Micrometer
ω	Omega

Amino Acid Abbreviations

A	Alanine
D	Aspartic acid
E	Glutamic acid
F	Phenylalanine
G	Glycine
H	Histidine
I	Isoleucine
L	Leucine
N	Asparagine
P	Proline
T	Threonine
W	Tryptophan
Y	Tyrosine

Chapter 1. Background

1.1. Enzymes in Biocatalysis

Enzymes are natural catalysts and preferred for industrial-scale catalysis over chemical catalysts.[1] This is due to their distinct advantages including:[2, 3]

- High efficiency: Enzyme catalysed reactions occur with an accelerated rate compared to those of the corresponding non-enzymatic reactions, usually $\frac{k_{cat}}{k_{uncat}}$= 10^5 to 10^{10}.[2, 3] This acceleration has been reported to a remarkable value of 10^{17} for orotidylate decarboxylase.[4]

- Environmentally friendly: Enzymes are degradable polypeptide making them environmentally benign biocatalysts.

- Mild reaction conditions: Most enzymes operate in a range of about pH 5-8, typically around 7, and in a temperature range of 20-40 ℃, preferably around 30 ℃. This reduces the undesired side-reactions often linked with traditional chemical synthesis such as decomposition, isomerization, racemization and rearrangement.

- Chemoselectivity: Enzymes are able to act on one type of functional group within a molecule without affecting further potentially reactive functionalities. This would offer a 'cleaner' reaction where the laborious purification steps to separate product(s) from the impurities can be avoided.

- Regioselectivity: Enzyme may distinguish between functional groups which are chemically located in different regions of the same substrate molecule. Therefore, enzymes are capable of transformations without the need for extra protection and deprotection steps that are usually required in organic synthesis.

- Enantioselectivity: Enzymes are chiral molecules since they are made from L-amino acids. Consequently, any type of chirality exist in the substrate is recognized upon formation of enzyme-substrate complex.
  Despite their advantageous, the application of enzymes in organic chemistry and industrial biotechnology have been prevented due to the following limitations:[5]

- Narrow substrate range when utilising nonnatural substrate.

- Poor stability under operating conditions: Enzymes are often applied in environments that differ significantly from their native environment such as organic solvents, high substrate and product concentrations and high salt concentrations, high temperatures and sub-optimal pH conditions.[6]

1.2. Carbon dioxide (CO_2) as a C1 Building Block

Carbon dioxide (CO_2) is one of the major greenhouse gases formed from fossil fuels. Since the start of the Industrial Revolution, the annual emission of CO_2 has increased more than 10-fold, corresponding to a 46% increase in the atmospheric CO_2 level, from about 280 ppm (parts per million) in the 1760s to 409 ppm today.[7] Although CO_2 is a critical factor for global warming, at the same time it is an important C1 carbon source for a future green chemistry. It is readily available, inexpensive and a less toxic carbon source compared to most of chemicals used in industrial processes.[8] It is also non-flammable which makes it easier for transportation.
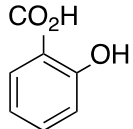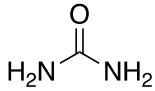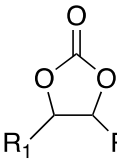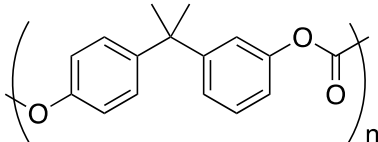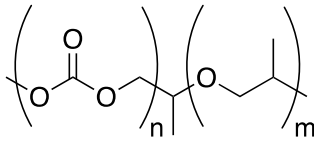
The development of CO_2-fixation reactions for the synthesis of organic molecules, which would enable transforming a waste gas into valuable chemicals, is one of the major challenges in synthetic organic chemistry.[9] One reason is that chemical conversion of CO_2 is considered energetically costly. CO_2 has the highest oxidation state of carbon (+4) and the lowest energy level, and its un-catalysed transformation by the chemical industry requires large energy input and/or harsh reaction conditions (Scheme **1-1**).[9, 10]
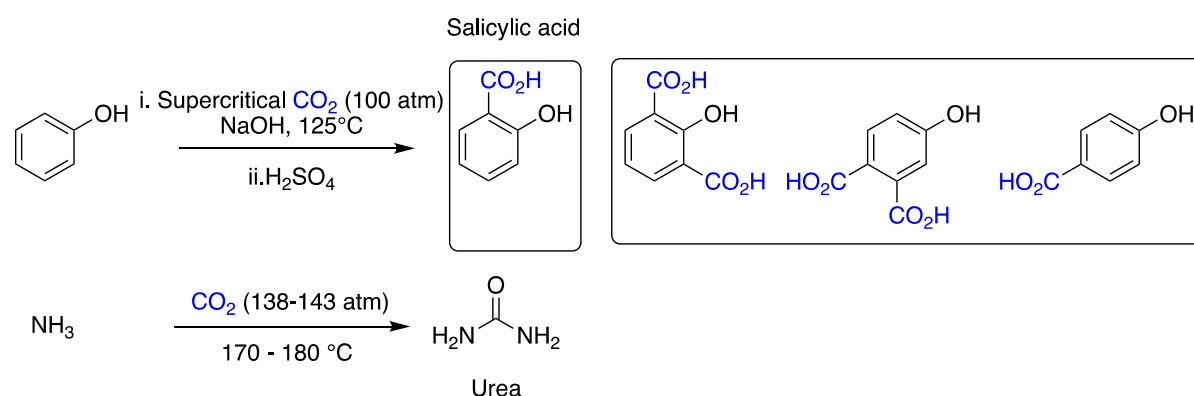
Scheme 1-1. Utilisation of CO_2 as carbon source in industrial processes.[10]

As a consequence, there are only a limited number of chemicals that are synthesised using CO_2 as raw material on industrial scale (Table 1-1).[11]

Table 1-1. List of major chemicals annually produced from CO_2.[11]

In nature, carbon dioxide is formed naturally from volcanic activity and the decomposition of biomass. On the other hand, nature plays an important role in the biological transformation of CO_2 in which 100 Gt/year of CO_2 is converted into organic compounds.[12] Nearly 95% of the CO_2 is transformed by CO_2-fixing enzymes, such as carboxylases, and only 5% through CO_2 reduction, highlighting the importance of carboxylases.[13]

Autotrophs are organisms that can grow using CO_2 as their only source of carbon.[9] In autotrophic organisms, complex organic substances such as carbohydrates, fatty acids and proteins are synthesised by assimilation of CO_2 through reduction.[14] Autotrophs are classified into two different groups based on the type of redox equivalents they use for reduction of CO_2: chemoautotrophs and photoautotrophs. Chemoautotrophs are organisms that obtain electrons required for CO_2 reduction from inorganic sources such as H_2, H_2S, elemental sulfur, metal ions, ammonia or nitrite. On the other hand, photoautotrophs derive electrons required for CO_2 reduction from water and use the energy of sunlight for CO_2-fixation.[9]

Nevertheless, all autotrophic organisms fix carbon through four main pathways:

1. Calvin-Benson-Bassham-cycle (Calvin cycle),[15]
2. Reductive tricarboxylic acid (TCA) cycle (Arnon-Buchanan cycle),[16]
3. Reductive Acetyl-CoA pathway (Wood-Ljungdahl pathway),[17]

4. Acyl-CoA carboxylase pathways: 3-hydroxypropionate/malonyl-CoA cycle,[18] 3-hydroxypropionate/4-Hydroxybutyrate cycle,[19] dicarboxylate/4-hydroxybutyrate pathway,[20] and the ethylmalonyl-CoA pathway.[21]

1.2.1. Calvin cycle with RuBisCO enzyme

The first autotrophic CO_2 fixation mechanism was elucidated by Calvin in 1961.[15] The Calvin cycle is found in photosynthetic organisms, mainly in plants on land and algae in water, and photosynthetic prokaryotes (cyanobacteria) which perform oxygenic photosynthesis. It is also found in autotrophic proteobacteria, some of which do not tolerate oxygen (anaerobic). In the first step of the Calvin cycle, CO_2 is fixed onto a pentose sugar by a key carboxylating enzyme called Ribulose-1,5-biphosphate carboxylase/oxygenase RubisCO (EC; 4.1.1.39).[22] The RubisCO enzyme is also found in some other bacteria and some archaea, but these lack other enzymes required for the cycle and/or there is no evidence for autotrophic growth.[23] As drawn in Scheme 1-2, the enzyme catalyses the reaction of CO_2 with ribulose-1,5-biphosphate (RuBP) , yielding two equivalents of 3-phosphoglycerate (3-PG). The carboxylation step catalysed by Rubisco involves a series of chemical reactions including:

- Enolisation and carboxylation: The substrate RuBP undergoes enolization which then acts as a nucleophile and attacks onto CO_2.
- Hydration: The β-keto acid intermediate generated from the previous step reacts with water and forms a geminal diol.
- C-C cleavage: The deprotonation of the geminal diol initiates the C-C cleavage which results in the formation of one molecule of 3-phosphoglycerate (3-PG) and a C-2 carbanion of 3-PG. The second molecule of 3-PG is generated by protonation of the C-2 carbanion of 3-PG.
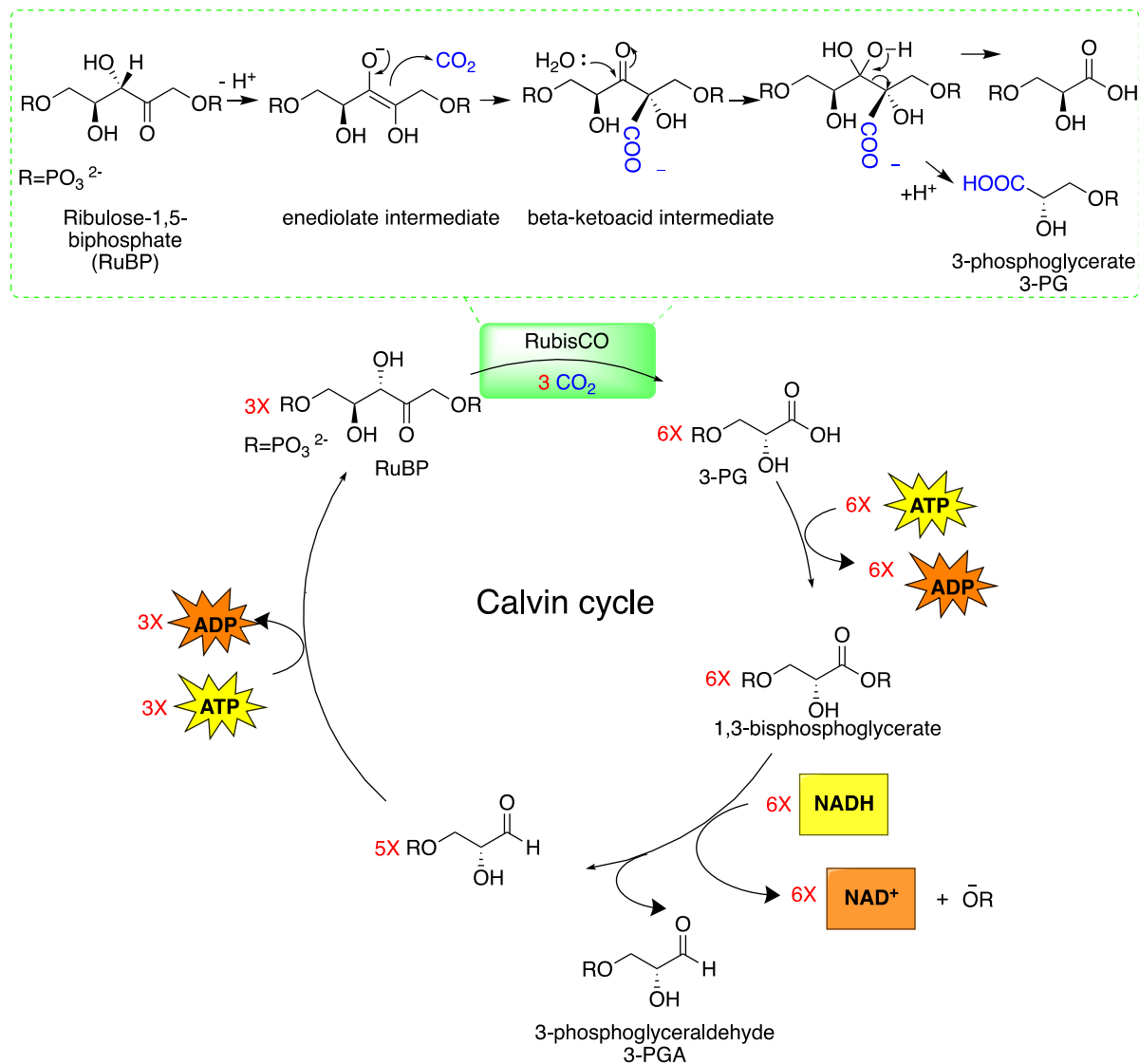
Scheme 1-2. Calvin Cycle, CO_2 fixation step catalysed by RubisCO highlighted in green[24]

Subsequently, 3-PG is phosphorylated with adenosine triphosphate (ATP) to produce 1,3-bisphosphoglycerate, which is then reduced to generate 3-phosphoglyceraldehyde (3-PGA). For every six equivalents of 3-PGA, five are used to regenerate the RuBP starting material. [24] Although RubisCO plays an important role in fixation of CO_2 in nature, its practical application as a biocatalyst is limited by its low turnover number of 5 s^{-1}.[25]

1.2.2. Reductive tricarboxylic acid (TCA) cycle (Arnon-Buchanan cycle)

The second autotrophic CO_2 fixation mechanism was first proposed by Evans *et al.* in 1966.[16] Evans *et al.* identified a green sulfur photosynthetic bacterium, *Chlorobium thiosulfatophilum*, which utilises a mechanism for autotrophic CO_2 fixation.[16] The mechanism for this bacterium

involves, CO_2 fixation reactions through a reductive tricarboxylic acid (TCA) cycle. In 1990, all the details of autotrophic CO_2 fixation through reductive tricarboxylic acid (TCA) cycle were worked out by Arnon and Buchanan, and therefore it is sometimes called the Arnon-Buchanan cycle.[26] The reductive TCA cycle is also found in archaea and several other groups of bacteria.[23]
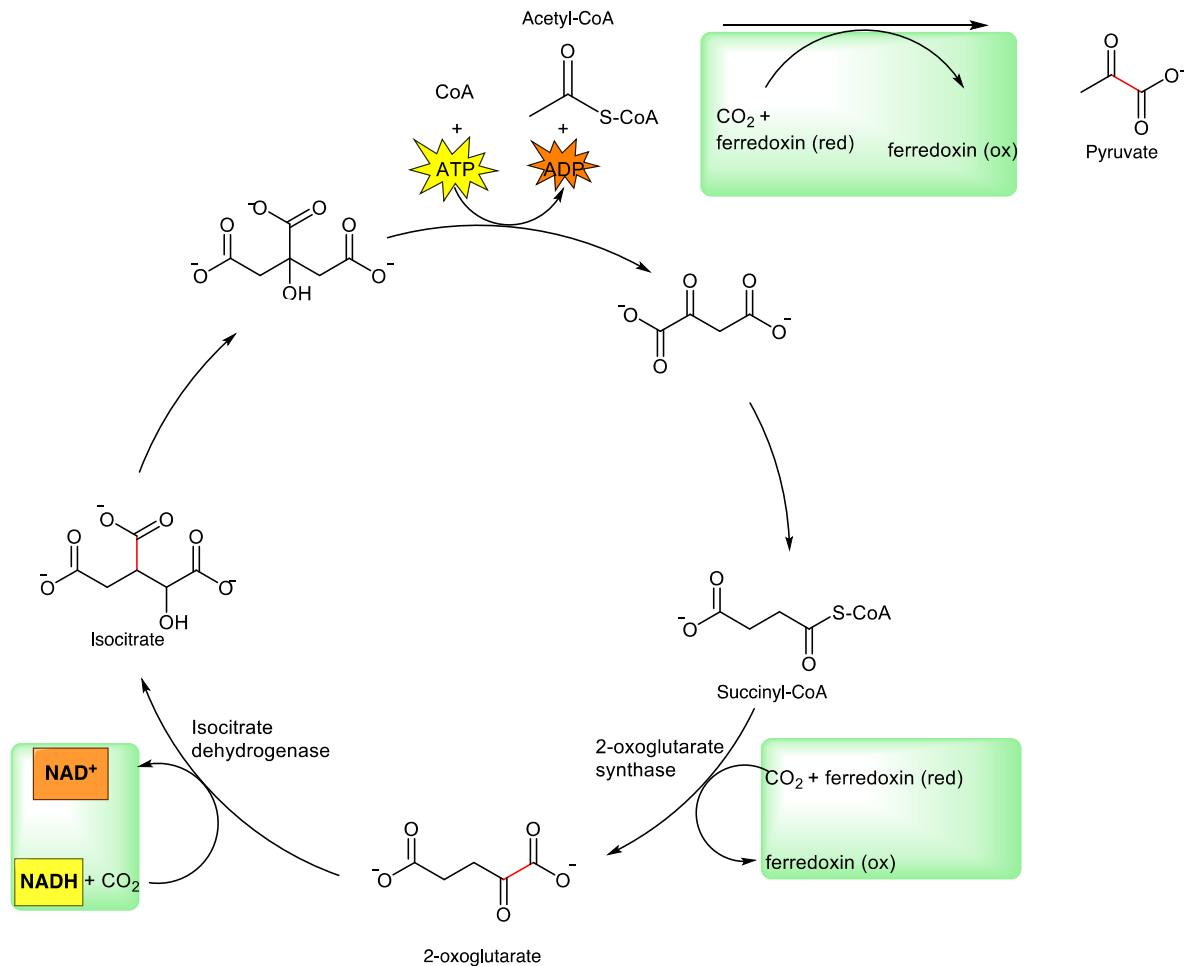
Scheme 1-3. Reductive TCA Cycle, three CO_2 fixation steps highlighted in green[24]

As drawn in Scheme 1-3 , there are three CO_2 fixation steps in the reductive TCA cycle two of which occur in a successive manner. First succinyl-CoA is reductively carboxylated by a ferredoxin dependent enzyme called 2-oxoglutarate synthase (EC; 1.2.7.3) to generate 2-oxoglutarate. Next, 2-oxoglutarate is also reductively carboxylated by isocitrate dehydrogenase (EC; 1.1.1.41) to form isocitrate. Finally, acetyl-CoA is reductively carboxylated by pyruvate synthase (EC; 1.2.7.1) to produce pyruvate.[24]

1.2.3. Reductive Acetyl-CoA pathway

The reductive acetyl-CoA pathway was first proposed by Ljungdahl and Wood in 1969, and hence is also referred as Wood-Ljungdahl pathway.[17]
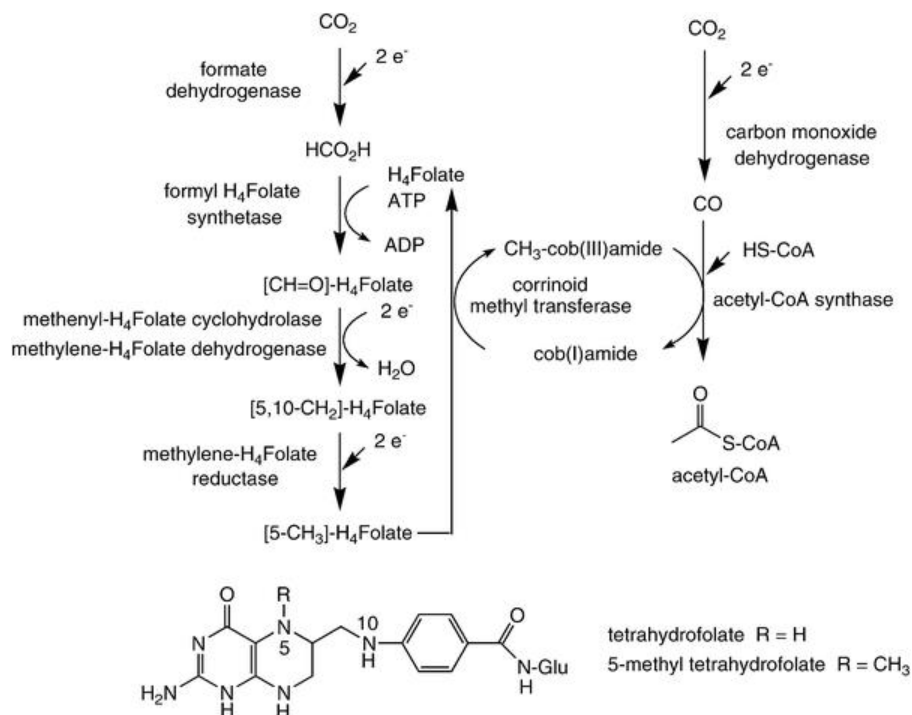
Scheme 1-4. Reductive acetyl-CoA pathway.[9]

CO_2 fixation through the reductive acetyl-CoA pathway operates in anaerobic organisms such as acetate forming bacteria and methane forming archaea.[23] The reductive acetyl-CoA pathway is a linear pathway which comprises two branches, methyl and the carbonyl branch. In this pathway two molecules of CO_2 are reductively coupled to ultimately form one unit of acetyl-CoA. As presented in Scheme 1-4, CO_2 is first reduced to formate by a nicotinamide adenine dinucleotide (NADH) dependent enzyme called formate dehydrogenase (FDH). The second molecule of CO_2 is also reduced to carbon monoxide by carbon monoxide dehydrogenase (CODH). Finally, the products from the two branches couple to yield one unit of acetyl-CoA.[24]

1.2.4. Acyl-CoA pathways

The fourth pathway of CO_2 fixation was discovered in a green non-sulfur bacterium, *Chloroflexus aurantiacus* in 2001.[18] A similar pathway was also discovered in an Archea,

Metallosphaera Sedula, in 2007.[19] In both pathways, acetyl-CoA serves as starting material, and carboxylation of acetyl-CoA and propionyl-CoA are the main CO_2 fixation reactions.
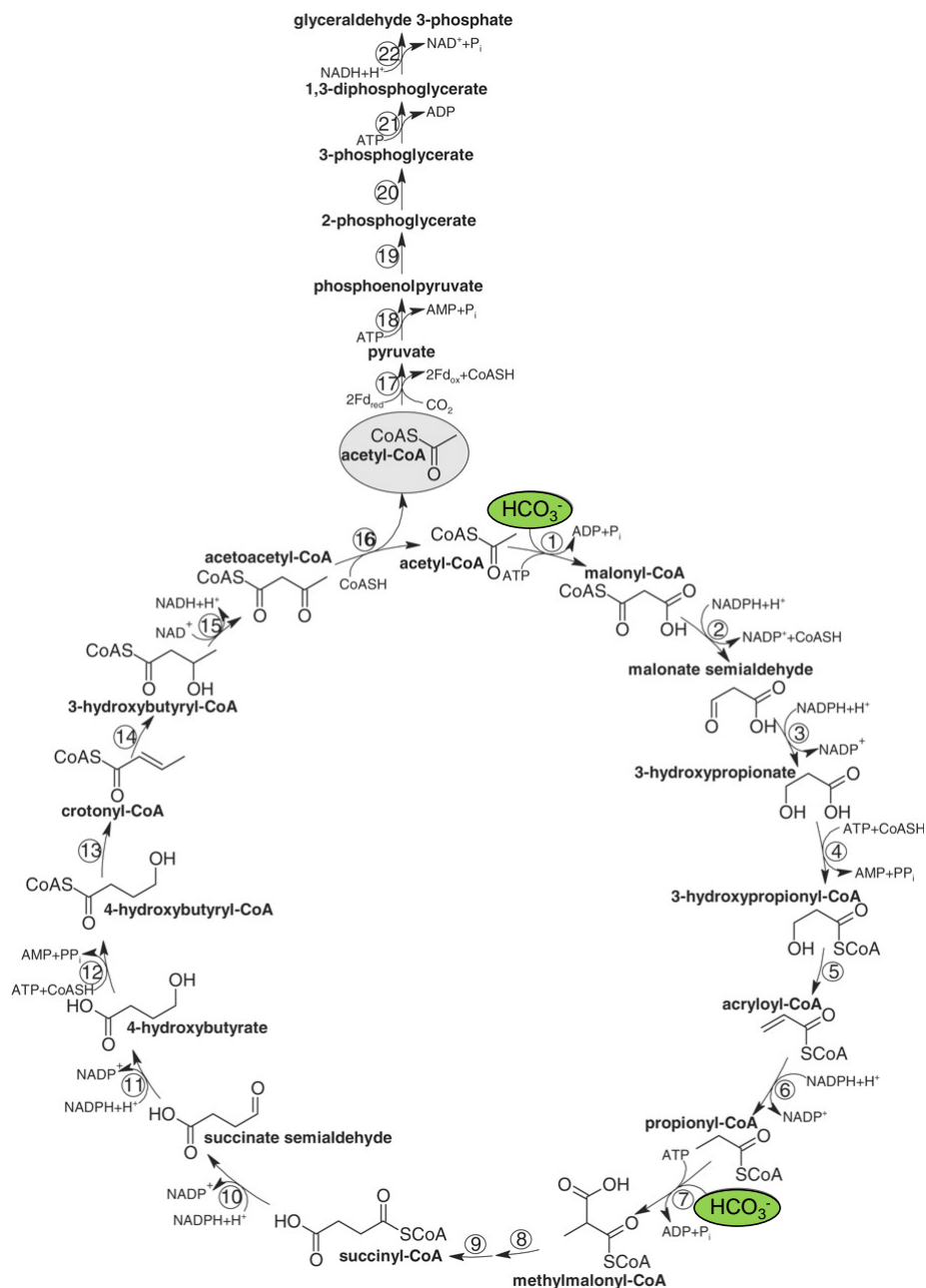
Scheme 1-5. The 3-hydroxypropionate/4-hydroxybutyrate cycle in *M.Sedula*, two CO_2 fixation steps are highlighted in green.[27]

As drawn in Scheme 1-5 , in the first CO_2 fixation reaction acetyl-CoA is carboxylated to malonyl-CoA by an ATP dependent enzyme called acetyl-CoA carboxylase using HCO_3^- as the carbon source. In the second carboxylation step, propionyl-CoA is carboxylated to methylmalonyl-CoA catalysed by propionyl-CoA carboxylase in a same fashion as the previous step.

Another form of the above CO_2 fixation reaction was discovered in a hyperthermophilic Archaeum, *Ignicoccus hospitali*.[20] The cycle is known as dicarboxylate/4-hydroxybutyrate cycle.[28] In this cycle, acetyl-CoA is carboxylated to pyruvate by a ferredoxin dependent enzyme called pyruvate synthase. The pyruvate is then phosphorylated with ATP to yield phosphoenolpyruvate. In the second CO_2 fixation reaction, phosphoenolpyruvate is carboxylated to oxaloacetate by phosphoenolpyruvate oxaloacetate lyase (Scheme 1-6).
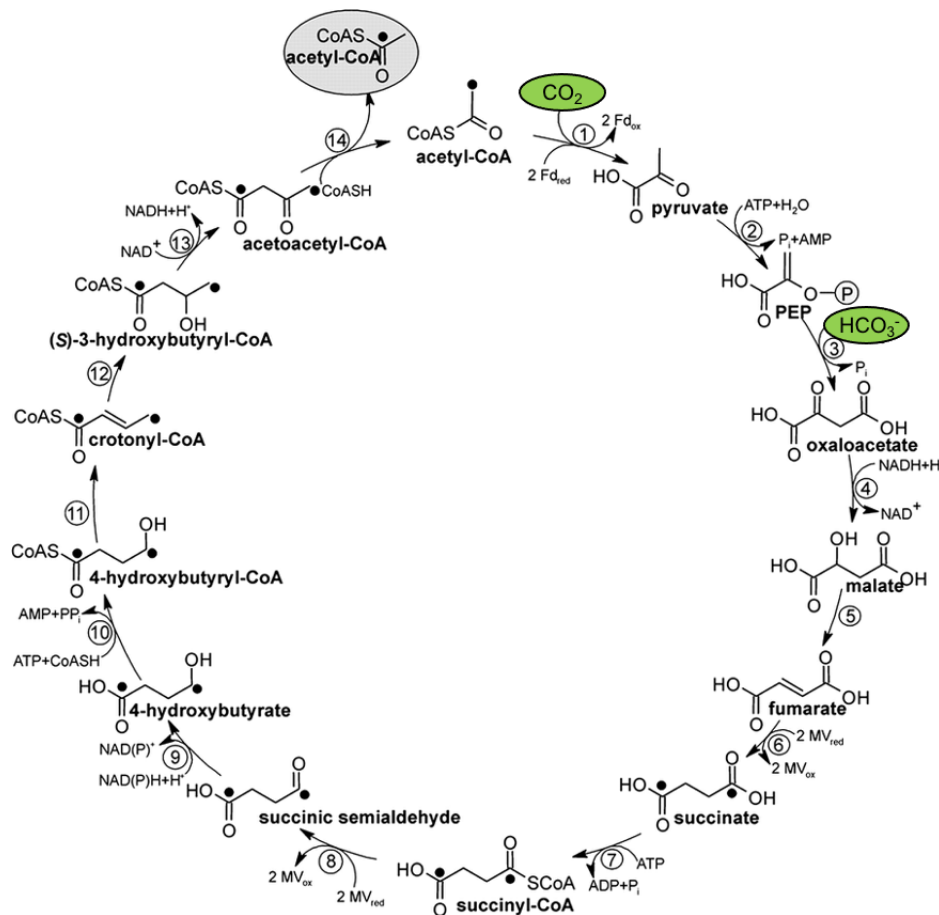
Scheme 1-6. The dicarboxylate/4-hydroxybutyrate cycle in *I.hospitali*, two CO_2 fixation steps are highlighted in green.[20]

1.3. Biocatalytic carboxylation

Generally, CO_2-fixing enzymes involved in natural metabolic pathways have been exploited in biotechnology for the transformation of CO_2, *via* either direct reduction of CO_2 or carboxylation of another substrate.[24] These enzymes naturally occur in the biosynthetic carboxylation pathways and catabolic (de)carboxylation pathways. Those involve in the biosynthetic carboxylation are highly specific. However, catabolic (de)carboxylase enzymes

have less substrate specificity and act on a wide range of substrates making them a better candidate for the biotechnological transformation. Catabolic (de)carboxylases have been utilised to a limited extent for the preparative biological transformation of CO_2, however the exact role of these enzymes and their 'true' substrates in nature are still unclear.[9] According to the Le Chatelier's principle, the carboxylation reaction may be achieved by pushing the position of the equilibrium towards the carboxylation by either the use of CO_2 in excess as concentrated carbonate/bicarbonate buffer at atmospheric pressure or in supercritical CO_2 (Scheme 1-7).[9] Alternatively, the reaction can be performed at a pressurized CO_2 environment in which the carboxylation reaction is coupled to an energetically more favourable reaction.[29]
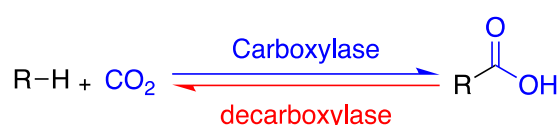
Scheme 1-7. Biocatalytic Carboxylation by de (carboxylases)

Enzymatic carboxylation reactions catalysed by catabolic (de)carboxylases with potential application in biocatalysis are categorised in three categorise based on their substrates:

- Epoxide substrate
- Aromatic and heteroaromatic substrate
- Aliphatic aldehyde substrate

1.3.1. Carboxylation of epoxide substrate

The enzymatic carboxylation of epoxide was discovered for the first time in *Xanthobacter* strain Py2 .[30, 31] In this organism, β-keto butyrate is generated from the direct fixation of CO_2 onto epoxypropane (propylene oxide) as part of epoxide degradation pathway (Scheme 1-8). The product of carboxylation is subsequently reduced to form β-hydroxy butyrate followed by polycondensation.[9]

Yield 7%
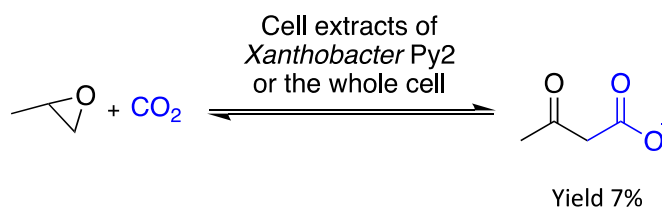
Scheme 1-8. Biocatalytic carboxylation of epoxypropane by *Xanthobacter* Py2

In this pathway coenzyme M (2-mercaptoethanesulfonate) attacks the epoxide resulting in a 2-hydroxy-alkyl sulfide catalysed by epoxyalkane CoM transferase. Then the oxidation of 2-hydroxy-alkyl sulfide by a dehydrogenase results in the formation a 2-oxoalkylsulfide which is the substrate for the carboxylation reaction catalysed by ketopropyl-CoM oxidoreductase/carboxylase (2-KPPC) (Scheme 1-9).[32]
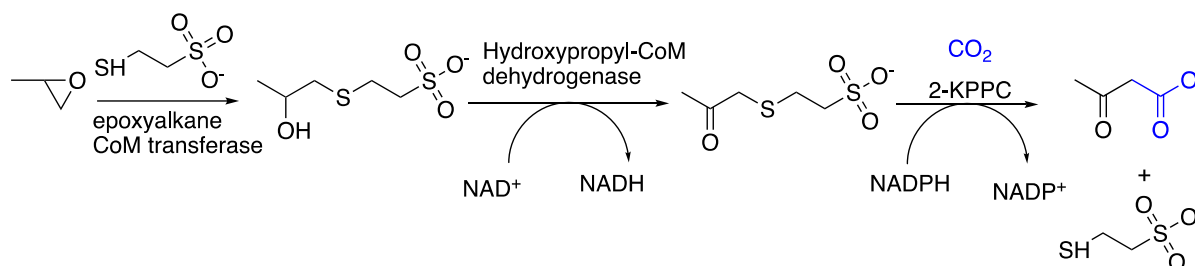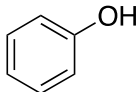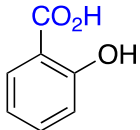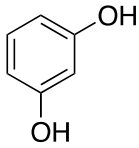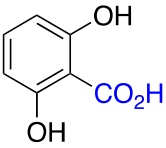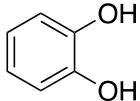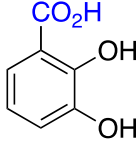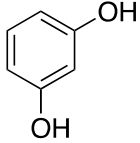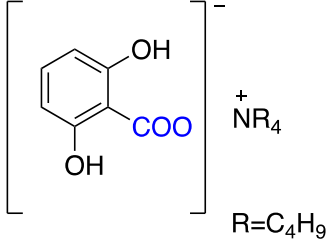
Scheme 1-9. Epoxide degradation pathway in *Xanthobacter* Py2.[32]

1.3.2. Carboxylation of aromatic and hetero-aromatic substrates

Aromatic non-oxidative decarboxylases have been widely employed in their reverse reaction for fixation of CO_2 (Table 1-2). For example, salicylic acid decarboxylase from *Trichosporon moniliiforme* was reported to carboxylate phenol in the presence of saturated concentration of $KHCO_3$ (3M). [33] This reaction which is the enzymatic version of the Kolbe-Schmitt reaction benefits from high level of regioselectivity as salicylic acid is the only product of the reaction. Also, the reaction is performed under mild temperature (30 ℃).

Table 1-2. Application of aromatic decarboxylases in CO_2 fixation

Substrate	Enzyme	product	Yield (%)	Ref
	Salicylic acid decarboxylase from *Trichosporon moniliiforme*		53	33
	2,6-dihydroxybenzoic acid decarboxylase from *Rhizobium* sp		48	34

	2,3- dihydroxy-benzoic acid decarboxylase from *Aspergillus oryzae*		43	35
	2,6-dihydroxybenzoic acid decarboxylase from *Rhizobium* sp	$R=C_4H_9$	Up to 95	36

In another examples, 2,6-dihydroxybenzoic acid decarboxylase from *Rhizobium* sp (2,6-DHBD-Rs),[34] and 2,3-dihydroxy-benzoic acid decarboxylase from *Aspergillus oryzae* (2,3-DHBD-Ao),[35] have been reported for their carboxylase activities. Very recently, the yield of carboxylation reaction catalysed by 2,6-DHBD-Rs was improved to up to 97% by adding quaternary ammonium salts into the reaction.[36] According to the Le Chatelier's principle, the equilibrium of the reaction was shifted towards carboxylation by precipitating the product of the reaction as an ammonium salt.[36]

Aromatic *N*-heterocycles, such as pyrroles and indoles, are also good substrates for enzymatic (de)carboxylation. For example, carboxylation of pyrrole was achieved regioselectively at the *ortho* position using the reverse reaction of pyrrole-2-carboxylate decarboxylase from *Bacillus megaterium* (Scheme 1-10).[37-40] The carboxylation reaction was favoured by using a saturated carbonate buffer thereby shifting the position of the equilibrium towards the carboxylation reaction.
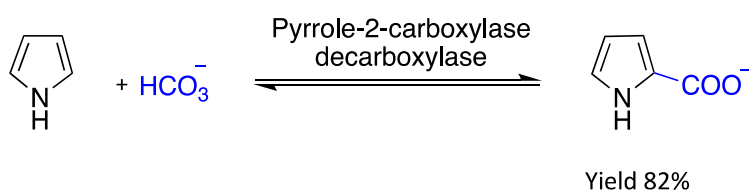
Scheme 1-10. Synthesis of pyrrole-2-carboxylate by Pyrrole-2-carboxylate decarboxylase from *B. megaterium*

Pyrrole-2-carboxylate decarboxylase from *Bacillus megaterium* is a homodimer and has a molecular mass of nearly 98 kDa. This enzyme has a unique feature as it requires an organic acid such as acetate, propionate or butyrate for its catalytic activity.[41]

The mechanism of action has been proposed for the above reaction, including the role of organic acid as a cofactor. The proposed mechanism uses an electrophilic substitution at the C2 carbon of pyrrole. It is suggested that the organic acid (cofactor) deprotonates at NH moiety, resulting in nucleophilic attack at the electrophile CO_2. After tautomerization, the intermediate regains aromaticity (Scheme 1-11).[39] This mechanism would need to be further investigated due to the low acidity of the NH of pyrrole which has pKa (16.5) much higher than of an organic acid (pKa 5).
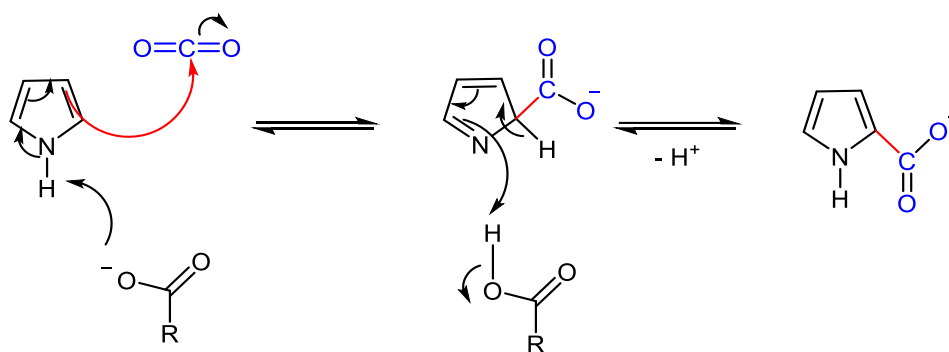
Scheme 1-11.Proposed reaction mechanism of pyrrole-2-carboxylate decarboxylase

In addition to pyrrole, the carboxylation reaction of indole was also achieved regioselectively at the 3-position using the reverse reaction of indole-3-carboxylate decarboxylase from *Arthrobacter nicotianae* and several moulds (Scheme 1-12). The reaction was observed by purified enzyme as well as by the resting cells. However, the yield for this reaction (34%) is not as high as the carboxylation reaction of pyrrole. It was assumed that the low yield is due to the low solubility of starting material, indole, in the reaction mixture.[42]
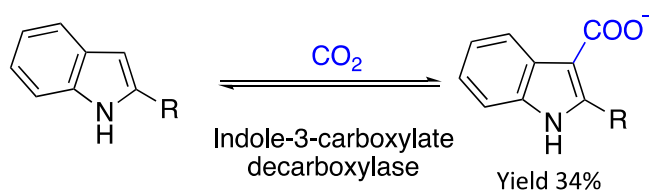
Scheme 1-12. Carboxylation of indole by indole-3-carboxylate decarboxylase

Recently, the enzymatic carboxylation of 2-furoic acid (FA) was also achieved to yield 2,5-furandicarboxylic acid (FDCA) using the reverse reaction of a decarboxylase enzyme from a thermophilic organism called *Pelotomaculum thermopropionicum* (Scheme 1-13). The enzyme is a member of prenylated flavin (prFMN) dependent family.[43]

Yield 4%
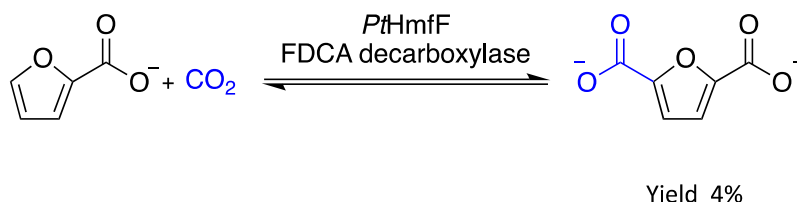
Scheme 1-13. Carboxylation of 2-furoic acid by *Pt*HmfF decarboxylase.

The reaction was tested using different sources of CO_2 to provide an insight into the prenylated-FMN (prFMN) dependent carboxylation (Table 1-3).[43] Accordingly, the maximum yield (4%) was observed in the presence of 1M $KHCO_3$ (pH 7.5). The treatment of the reaction mixture with pressurized CO_2 (32 bar) had no significant improvement on the yield of the reaction showing the importance of the bicarbonate at (1M) to achieve the carboxylation reaction at . Performing the reaction in the absence of bicarbonate only resulted in 0.3% FDCA under 32 bar CO_2, whereas no FDCA was produced under N_2. The future work to improve the yield of the reaction was reported to be *in situ* conversion of FDCA to overcome the unfavourable equilibrium for the carboxylation reaction.

Table 1-3. Different sources of CO_2 applied to achieve the reverse reaction of *Pt*HmfF decarboxylase

Reaction condition	1M $KHCO_3$, N_2 at atmospheric pressure	1M $KHCO_3$, CO_2 (32 bar)	0.1 M KPi, N_2 at atmospheric pressure	0.1 M KPi, CO_2 (32 bar)
(%) Yield	~4	~4	0	~0.3

1.3.3. Carboxylation of aliphatic substrates

Pyruvate decarboxylase (PDC; EC 4.1.1.1) is a key enzyme in alcoholic fermentation and is perhaps the simplest thiamine diphosphate (pyrophosphate) (ThDP) dependent enzyme. The enzyme catalyses the decarboxylation of pyruvic acid to acetaldehyde and CO_2. In yeast, the

product of glycolysis, pyruvate, is cleaved to acetaldehyde first, after which the acetaldehyde is converted to ethanol by alcohol dehydrogenase (ADH; EC 1.1.1.1) (Scheme 1-14).[44]
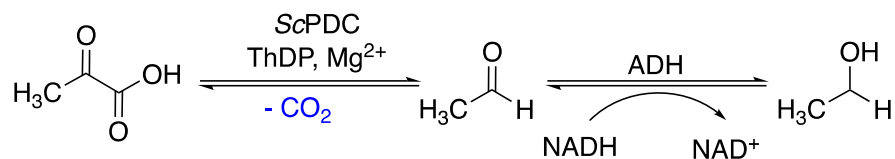
Scheme 1-14. Biocatalytic decarboxylation of pyruvic acid by PDC followed by a reduction step by ADH

The reverse reaction of this enzyme has attracted attention as a catalytic approach for carboxylation of aliphatic substrates. The enzyme has been reported for the reverse carboxylation reaction of acetaldehyde to produce pyruvic acid (Scheme 1-15).[45] The reaction was performed using PDC from *Saccharomyces cerevisiae* (*Sc*PDC) and ThDP as the cofactor. The position of the equilibrium was described to be directed towards the carboxylation reaction by using excess CO_2 in the form of a mixture of sodium carbonate/bicarbonate buffer. It was explained that the buffer system not only pushes the reaction towards the carboxylation, but also serves as a solvent. The highest conversion was reported at pH 11. Also, the effect of ionic strength of the carbonate buffer was tested with the maximum reported yield of reaction (81 %) using carbonate buffer at the concentration of 500 mM.[45]
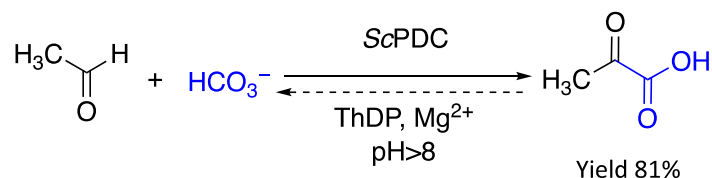
Scheme 1-15. Synthesis of pyruvic acid by *Sc*PDC

In 2002, a one-pot two steps enzymatic system was introduced for the production of the L-lactic from acetaldehyde and CO_2.[46] This example was the first ever reported example in which the reaction of a decarboxylase enzyme was assisted in its reverse direction by coupling it to a second reaction. The reaction comprised of two steps: carboxylation catalysed by *Sc*PDC and hydrogenation catalysed by a L-lactate dehydrogenase (L-LDH) (Scheme 1-16). Following the earlier publication,[45] the reaction intermediate, pyruvate, was synthesised by reversing the reaction of *Sc*PDC in the carbonate buffer, which then was reduced to L-lactic

acid in the presence of NADH. The maximum yield of (65%) was reported at pH 9.5 and the concentration of 500 mM of the carbonated buffer.[46]
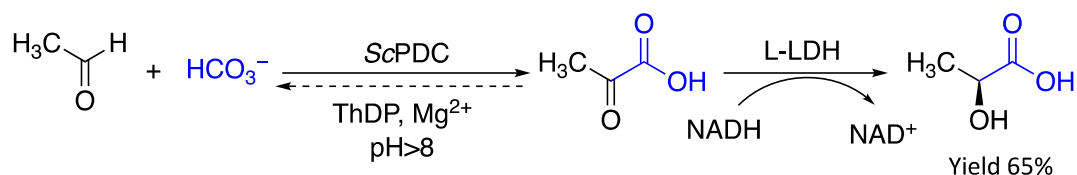
Scheme 1-16. One-pot enzymatic synthesis of L-lactic acid from acetaldehyde and carbonate buffer

Following the above example, in 2010 a triad enzymatic system was proposed for the synthesis of L-lactic to challenge the practicality of the CO_2 fixation reaction by a decarboxylase enzyme.[47] However, this system was slightly different compared to the previous example[46], since it involved an additional step prior to the carboxylation step, in which acetaldehyde was generated *in situ* from the oxidation of ethanol catalysed by an alcohol dehydrogenase enzyme, ADH (Scheme 1-17).
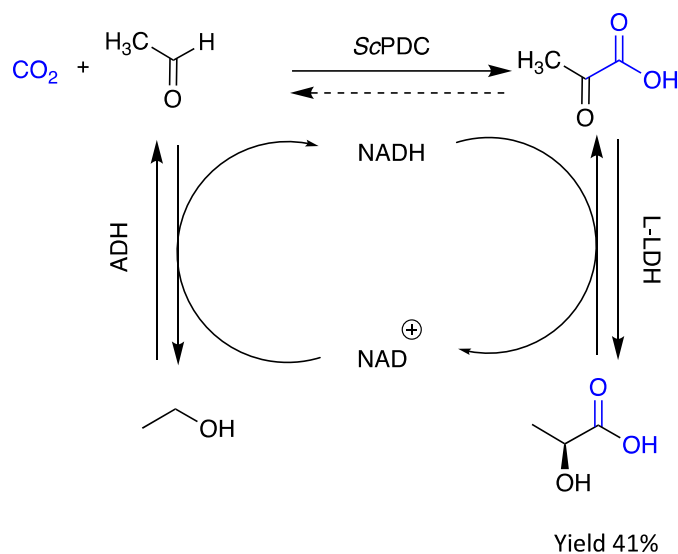
Scheme 1-17. The triad enzymatic reaction for the synthesis of L-lactic acid from ethanol and CO_2.[47]

The incorporation of an extra step was beneficial for the cascade since it simultaneously provided a cofactor regeneration system. This design not only helped to regenerate the cofactor but also minimised the handling of a toxic molecule such as acetaldehyde. Like the previous examples, carbonate buffer was used as a source of CO_2. But it is worth mentioning that this was not the only source of carbon as the solution was also purged with CO_2 gas. This

was claimed to be essential to provide the system with a saturated concentration of carbon source throughout the process. The maximum yield of 41% was reported at pH 9.5 and 250 mM of the carbonate buffer.[47]

Very recently, enzymatic carboxylation by a decarboxylase in conjugation with a transaminase reaction has been utilised as a biocatalytic approach in the stereospecific synthesise of α-amino acids from aldehydes.[29]
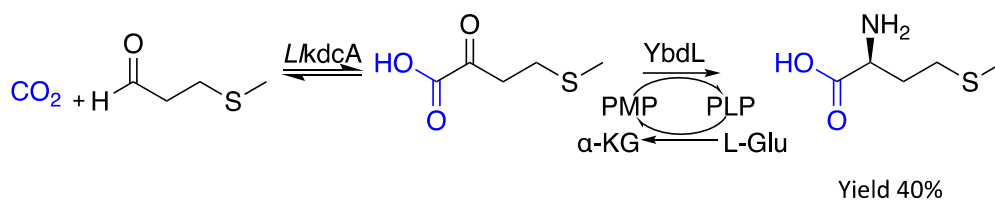
Scheme 1-18. Biocatalytic synthesis of L-Met from methional by *Ll*KdcA and *YbdL*

In this example, carboxylation of methional by a branched chain decarboxylase KdcA from *Lactococcus lactis* is coupled to a methionine aminotransferase (YbdL) from *E. coli K12* for the synthesis of L-methionine (Scheme 1-18). The reaction was performed in a pressurized CO_2 environment (2-8 bar). It was reported that L-methionine was produced with a yield of about 3%. The reaction was optimised by increasing the concentration of the first enzyme (*Ll*KdcA) and methional (from 1 mM to 4 mM) as well as reaction time (48 hours) to achieve a yield of 40%.[29]

1.4. Aims and objectives

From the introduction, it is clear that the reverse reaction of decarboxylase enzymes for the fixation of CO_2 is a challenging reaction, particularly with aliphatic substrates. The aim of this thesis is to investigate the application of pyruvate decarboxylase from yeast, *Saccharomyces cerevisiae* (*Sc*PDC) for the fixation of CO_2 on its own or in an artificial enzyme cascade to synthesis amino acids from aldehydes and CO_2. The cascade consists of one-pot two steps: carboxylation and transamination catalysed by *S. cerevisiae* L-alanine transaminase (*Sc*ALT). Also, the proposed cascade will be tested with *Ll*KdcA as the decarboxylase enzyme. The tasks required to achieve the overall aim of the thesis lead to following specific objectives:

- Cloning, expression and purification of the enzymes required for the cascade (*Sc*PDC, *Ll*KdcA and *Sc*ALT1)

- Testing the activity of both ScPDC and LlKdcA enzymes in decarboxylation and carboxylation reactions. Also, evaluate the effect of different sources of CO_2 on the carboxylase activity of above enzymes.
- Study the proposed cascade for the synthesis of amino acids (L-alanine and L-homoalanine).
- Investigate the parameters that influence the yield of the cascade (pH, pressure and substrate stability).
- Increasing the substrate spectrum of ScPDC for both carboxylation and decarboxylation reaction.
- Immobilisation of ScPDC with the potential for application in flow system for the proposed cascade.

Chapter 2. Materials and methods

All reagents and chemicals were of highest available purity and purchased from Merck, Thermo Fisher Scientific and New England BioLabs (NEB). [1]H nuclear magnetic resonance (NMR) spectra were recorded on a Bruker 400 MHz spectrometer. The chemical shifts (δ) were recorded in parts per million (ppm).

2.1. Preparation of *S. cerevisiae* genomic DNA

The genomic DNA of *S. cerevisiae* was extracted by the glass-bead method.[48] This method utilises disruption of the yeast cell wall by physical fractionation of cells with small glass beads. Following this method, cells were grown overnight in YPD medium (5 ml) at 30 °C in a roller drum at 600 rpm. The cultures were transferred to a falcon tube and centrifuged at $4000 \times g$ for 5 minutes. After removing the supernatant, the cells were washed with sterile water (3 ml) and centrifuged again as before. Cells were re-suspended in Lysis buffer (500 µl). Acid washed glass beads were added to the cells (up to 1.25 ml), and vortexed at maximum speed for 2 minutes and centrifuged at $3000 \times g$ for another 2 minutes. The liquid phase was removed and transferred to an Eppendorf tube. Ammonium acetate (275 µl, 7 M, pH=7) was added and incubated for 5 min at 65 °C and 5 min on ice. Subsequently, chloroform (500 µl) was added, vortexed and centrifuged at $3000 \times g$ for 30 seconds. The top layer was collected and incubated with isopropanol (1ml) for 5 minutes at room temperature. The liquid was centrifuged at $3000 \times g$ for 5 minutes. The obtained pellets were washed with 70% ethanol (500 µl) and centrifuged at $3000 \times g$ for 5 minutes. The pellets were dissolved in sterile water (20 µl).

YPD medium: Yeast extract (10 g), Peptone (20 g) and Dextrose (20 g) in 1L of distilled water. The YPD medium was autoclaved prior to use.

Lysis buffer: 0.1 M Tris-HCl (pH 8.0), 50 mM EDTA, 1% SDS

2.2. DNA manipulation

PCR was used to amplify target gene for cloning into a vector. All primers were purchased from Merck. For amplification of a gene fragment, an appropriate elongation time and increased number of cycles were applied.

Following the manufacture's protocol, a PCR reaction was set up in a PCR tube containing the reagents shown in (Table 2-1). The Q5 High Fidelity DNA polymerase within the Hot start High-Fidelity Master Mix amplifies the gene of interest with the highest fidelity ($\sim$280 times) compared to Taq polymerase.

Table 2-1. Reagents and their corresponding concentrations used in a PCR reaction

Component	25 µL Reaction	Final Concentration
Q5 Hot Start High-Fidelity 2× Master Mix	12.5 µL	1 ×
10 µM Forward Primer	Variable (usually 1.25 µL)	Variable
10 µM Reverse Primer	Variable (usually 1.25 µL)	Variable
Template DNA (15 ng/µL)	1 µL	0.6 ng/µL
Nuclease-free water	To 25 µL	

The PCR tube was transferred to a PCR thermocycler machine and the PCR reaction was performed following the condition shown in (Table 2-2)

Table 2-2. PCR reaction set up

Step	Temperature	Time
Initial Denaturation	98 °C	30 seconds
35 cycles	98 °C	10 seconds
(Denaturation-Annealing-Elongation)	50-72 °C	30 seconds
	72 °C	30 seconds per kb
Final Extension	72 °C	2 minutes
Hold	4 °C	

The DNA template was denatured at 98 °C for 30 seconds, and then the temperature was lowered to an appropriate annealing temperature (T_a) for 30 seconds allowing the PCR primers to bind with their complementary strand at the denatured template DNA. Primers were extended by DNA polymerase at 72 °C. The extension time was varied depending on the size of target DNA (30 seconds per kb).

The list of primers used within this thesis are summarised in (Table 2-3).

Table 2-3.List of primers

Name	Sequence 5' to 3'
PDC1-FWD (BamHI site underlined)	CGCCG<u>GGATCC</u>TCTGAAATTACTTTGG
PDC1-REV (XhoI site underlined)	CGG<u>CTCGAG</u>TTATTGCTTAGCGTTGGTAG
pJ414-backbone FWD	CTCGAGCCCCCTAGCATAAC
pJ414-backbone REV	GGATCCACGCGGCACCAG
*Sc*ALT1 FWD	aactggtgccgcgtggatccATGTTATCACTGTCTGCC
*Sc*ALT1 REV	gttatgctaggggggctcgagTCAGTCACGGTATTGGTC
*Sc*ALT1-MSP-FWD	CAATCTTCGCTAAACGACCTGCG
*Sc*ALT1-MSP-REV	GGATCCACGCGGCACCAG
pJexpress414Rev primers	CTCAAGACCCGTTTAGAGGC
T7F	TAATACGACTCACTATAGGG

2.2.1. Agarose gel electrophoresis of DNA

Gel electrophoresis was used to separate and detect DNA. Agarose DNA electrophoresis gels (1.2% w/v) were prepared by dissolving agarose (0.6 g) in 50 mL of 1× TAE buffer (40 mM Tris, 1 mM EDTA, pH 8.0) and heating in a microwave. The solution was cooled on bench before the addition of 1µL of ethidium bromide from (10 mg/ml) stock solution. The agarose gels were then cast in a gel tank with well-gel comb and allowed to set. Gel electrophoresis was performed in a tank containing 1×TAE as running buffer at 100 V for 40 minutes. DNA fragments were compared against 100 bp or 1 Kb DNA ladder with known fragments size. Gels were imaged under UV light.

2.2.2. Restriction enzyme digestion of DNA

Restriction enzymes, BamHI-HF and XhoI were purchased from NEB and digestions were performed according to the conditions given by NEBcloner tool under digestion subsection. For example, for insertion of *Sc*PDC fragment into pJexpress414 vector double digestion protocol was performed using both restriction enzymes (BamHI-HF and XhoI) in one reaction. As a general protocol, 1 µg of DNA was mixed with 1 µL of each enzyme, 5 µL of 10× CutSmart

Buffer made up to a final volume of 50 µL with nuclease free water. The samples were incubated at 37 °C for 1 hour. After incubation period, 4 µL of 6× gel loading buffer was added to each reaction. Digested DNA fragment was analysed by 1.2 % agarose gel electrophoresis. The DNA fragment was excised from the gel using a sharp scalpel. The DNA fragment was purified using QIAquick gel extraction kit.

CutSmart buffer contains: 50mM potassium acetate, 20mM Tris-acetate, 10mM Magnesium acetate,100 µg/ml BSA (pH 7.9 at 25 °C)

2.2.3. Ligation of DNA

For the ligation of digested vector DNA (pJexpress414) and target insert DNA ligation protocol with T4 DNA Ligase (NEW ENGLAND BioLabs) was followed. A total of 50 ng of DNA (with a molar ratio of insert to vector 3:1) was mixed with 2 µL T4 DNA ligase buffer, 20 µL nuclease free water and 1 µL T4 DNA ligase. The samples were mixed with gentle pipetting, centrifuged briefly and incubated at 16 °C overnight. The ligated DNA fragment was transformed into the *E. coli* cells (DH5α) following the heat-shock protocol.

2.2.4. Extraction of plasmid DNA from E. coli

Bacterial colonies were picked and grown in a 10 mL overnight culture. The antibiotic, ampicillin, was added to maintain plasmid selection. Plasmids were isolated using Wizard® Plus SV Minipreps DNA Purification System from Promega according to the centrifugation protocol. Cells from 10 mL overnight culture were harvested by centrifugation at 4000 rpm for 10 min. The cells were re-suspended, lysed and neutralized using the buffers and their appropriate volume according to Table 2-4.

Table 2-4. Buffers used in plasmid isolation

Solutions	Volume (µL)
Cell resuspension	250
Cell Lysis	250
Alkaline Protease	10
Neutralization	350

Cell debris and chromosomal DNA were removed by centrifugation at 14000 rpm for 10 min. Then, the plasmid DNA was selectively bounden to a silica column. This was achieved by applying the cleared lysate to the spin column and centrifugation at 14000 rpm for 1 min. The flow through was discarded and the column was reinserted into the collection tube. The silica column was washed with wash solution to remove contaminants and dried by centrifugation at 14000 rpm for 1min. The flow through was discarded and an extra step of washing was performed at 14000 rpm 2 min to ensure ethanol removal. Finally, the plasmid was eluted with nuclease free water by adding water to the silica column followed by centrifugation at 14000 rpm for 1 min. The purified plasmid was stored at -20 °C.

2.2.5. Quantification of DNA

The concentration and purity of plasmid DNA or PCR product was measured using NanoDrop ND-1000 UV spectrophotometer. The absorbance of DNA was measured at 260 nm and 280 nm and A_{260}/A_{280} was used as an indicator for the purity of DNA. A A_{260}/A_{280} ratio of ~1.8 was accepted as pure DNA.

2.3. Growth of E. coli strains

E. coli cells (BL21 (DE3), DH5α and Rosetta) were commonly cultured in Luria-Bertani media (LB) with shaking (180 rpm) and on solid LB plate supplemented with agar at 1.5% (w/v). In order to maintain plasmid selection appropriate antibiotics were added to the media. Ampicillin was mainly used as the antibiotic in this work. To prepare LB media 25 grams of LB (Containing tryptone 10 g, NaCl 10 g and yeast extract 5 g) was added to 1 L of deionised water and autoclaved. To prepare LB agar, 7g of LB agar (containing tryptone 10 g, NaCl 10 g, yeast extract 5 g and Micro Agar 10 g) was added to 200 mL of deionised water and autoclaved.

2.3.1. Preparation of chemically competent E. coli

Within this thesis, BL21 (DE3) and DH5α strains were made chemically competent following the procedure described below. A glycerol stock of BL21(DE3) strain was obtained from Dr Andy Brennan, University of Nottingham, UK. For DH5α strain a stock of NEB 5-alpha purchased from NEB was used. Overnight cultures were set up by inoculating 10 mL LB media (without any antibiotics) with appropriate strain from their corresponding stock solutions. On

the following day, the starter culture was diluted with fresh LB media (1/100 to a final volume of 10 mL) and incubated at 37 °C with shaking (180 rpm) for 3 hours. In order to minimise any chance of contamination at this step OD_{595} measurements were excluded. The culture was then chilled on ice for 10 minutes and then centrifuged at $3000 \times g$ for 10 minutes at 4 °C in a 50 mL polypropylene tube. The supernatant was discarded, and pelleted bacterial cells were resuspended in 5 mL ice-cold sterile 50 mM $CaCl_2$ and incubated on ice for 1 hour. The cells were then harvested as above, and the pellets were resuspended in 600 μL of ice-cold 50 mM $CaCl_2$. The cells were aliquoted as 50 μL aliquots in sterile 1.5 mL Eppendorfs followed by addition of 10 μL of glycerol from 50% stock solution. Finally, aliquots were stored at − 80 °C until they required.

2.4. Heat-shock transformation of competent E. coli

For transformation, a vial of chemically competent cells (50 µL) was thawed on ice. Subsequently, plasmid DNA (2 µL, 100 ng/ µL) was added to the vial and this mixture was incubated on ice for 30 minutes. This step was followed by a thermal shock for 30 seconds at 42 °C, which opens the pores of cell membrane to introduce of the DNA plasmid. The mixture was immediately transferred on ice for 2 min, which closes the pores of cell membrane. Then 950 µL of SOC medium (2% Trypton, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$ and 20mM glucose) was added to the transformed cells and cells were grown at 37 °C with shaking (180 rpm). 40 µl of culture was transferred on an agar plate supplemented with appropriate antibiotic and incubate at 37 °C overnight.

2.5. Recombinant expression in E. coli

In this thesis, expression vectors were constructed for *Sc*PDC and *Sc*ALT1 by amplifying the corresponding genes from *S. cerevisiae* genomic DNA and sub-cloning into the pJexpress414 expression vector. *Ll*KdcA (pET32a) expression vector was purchased from DC Bioscience Ltd, UK. All three expression constructs were initially used to transform BL21(DE3) *E. coli* strain. *Sc*ALT1 expression vector was also used to transform Rosetta (DE3) *E. coli* strain. In order to evaluate the effect of different parameters on the expression levels and the solubility of the recombinant proteins, small-scale expression experiments were initially performed.

Table 2-5. List of genes used in this thesis

*Sc*PDC	Pyruvate decarboxylase from *Saccharomyces cerevisiae*
*Sc*ALT	Alanine transaminase from *Saccharomyces cerevisiae*
*Ll*KdcA	branched chain decarboxylase (KdcA) from *Lactococcus lactis*

2.5.1. Small-scale protein expression in E. coli (20 mL)

After transformation, single colonies were picked from LB agar plates and used to inoculate 10 mL of growth medium (LB) supplemented with ampicillin (0.1 mg.mL^{-1}) and cultured overnight. The next day, the starter culture was diluted with fresh LB medium (5/50 to a final volume of 50 mL in 250mL baffled flask) and incubated at 37 °C with shaking (180 rpm) until OD$_{595}$ of 0.6 was reached. The flask content was split into two smaller flasks (2×20 mL). The contents of one flask were treated as non-induced and the other as induced. Protein expression was induced by addition of IPTG (1 M stock in deionised water and sterile filtered) to a final concentration of 1 mM directly to the cell suspension. Depending on the experiment, induction temperature and incubation time were varied. After certain incubation period, cells were harvested by centrifugation. Supernatants were discarded and bacterial cell pellets were stored at -20 °C until protein extraction and analysis by SDS-PAGE for the presence of over-expressed protein.

2.5.2. Large-scale expression in E. coli (1L)

For large-scale expression in *E. coli*, single colonies were picked up and used to inoculate a 10 mL overnight culture (LB medium) supplemented ampicillin (0.1 mg.mL^{-1}). The next day, the starter culture was diluted with fresh LB medium (10/1000 to a final volume of 1L) containing appropriate antibiotic. The 1L culture was then grown at 37 °C with shaking (180 rpm) until OD$_{595}$ of 0.6 was reached. Recombinant protein expression was induced with IPTG (1M stock solution in sterile water) and cells were harvested.

2.5.3. Disruption of bacterial cell walls

The bacterial cell pellets previously stored in a 50 ml polypropylene tube was taken out and thawed at room temperature. The thawed cells were resuspended in appropriate buffer (usually 4 mL of buffer A for 1 g of wet cells), Table 2-6. The cells were mixed until

homogenous suspension obtained. The sonicator probe was immersed in half of the volume of resuspended cells while keeping it on ice. Pulses of 10 µm amplitude were applied to the sample in cycles of 30 s, followed by 30 s of no sonication to cool the sample for the total time of 5 minutes. After the first 5 min round, the process was changed to pulses of 15 µm amplitude in cycles of 30 s, followed by 30 s of no sonication for another 5 min.

Total cell lysate was transferred into a small centrifuge tube and centrifuged at $35000 \times g$ at 4 °C for 35 minutes to remove cell debris and obtain a soluble fraction. After centrifugation, the soluble fraction was further clarified by filtering through a 0.45 µm filter.

2.5.4. Standard immobilized metal (Ni^{2+}) affinity chromatography (IMAC)

Following expression recombinant proteins contained an N-terminal hexa (6×) His-tag, the AKTA purifier chromatography system was used for the IMAC purification, all buffer (Table 2-6) used for purification were degassed and filtered prior to purification. The 5 ml HiTrap chelating HP affinity column was used for the purification which was previously charged. The column was washed with 25 ml (5× column volume) degassed water and equilibrated with 25 mL of IMAC binding buffer (Buffer A). The soluble protein fraction was loaded to the column. The absorbance of protein was monitored at 280 nm. Once the absorbance started to rise the flow through was collected. The IMAC binding buffer was kept passed through the column until the absorbance dropped to the base line. Then the concentration of imidazole was increased from 0-1 M. The eluted fractions were collected and analysed by SDS-PAGE and Coomassie straining.

Table 2-6. Summary of the buffer compositions used in purification of ScPDC

Buffer Compositions	Buffer A (Lysis/Binding Buffer)	Buffer B (Elution Buffer)	Buffer C (Dialysis Buffer)
*Citrate buffer	50 mM	50 mM	50
NaCl	500 mM	500 mM	-
Imidazole	15 mM	700 mM	-
ThDP	-	-	1 mM
MgCl$_2$	-	-	2 mM
Glycerol	-	-	40%
pH	6.5	6.5	6.5

*Citrate buffer: Dissolve 9.6 g of citric acid in 800 mL of deionized H$_2$O. Adjust to pH 6.5 with 5 M NaOH. Adjust to a final volume of 1000 mL with deionized H$_2$O.

For LlKdcA after dialysis the sample was concentrated to 11 mg/ml using a vivaspin column (100 kDa MWCO). Finally, cofactors (MgCl$_2$ and ThDP) were added to final concentrations of 1mM and 0.5 mM respectively prior to the assay. The (His)$_6$-LlKdcA recombinant purified protein was stored at – 80 °C until needed.

Table 2-7. Summary of the buffer compositions used in purification of ScALT1

Buffer Compositions	Buffer D (Lysis/Binding Buffer)	Buffer E (Elution Buffer)	Buffer F
Sodium phosphate (Na$_2$HPO$_4$-NaH$_2$PO$_4$)	50 mM	50 mM	50 mM
NaCl	500 mM	500 mM	-
Imidazole	15 mM	1000 mM	-
PLP	-	-	1mM
Glycerol	-	-	10%
pH	8.0	8.0	8.0

2.5.5. Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

Both stacking and resolving gels were prepared using the ingredients summarized in Table 2-8. The 12% (w/v) resolving gel was prepared and placed between the two glasses of the Bio Rad assembly system. After 30 minutes resolving gel was set. The 5% (w/v) stacking gel was prepared and poured on top of the resolving gel. The 1 mm comb was placed on to the stacking gel immediately giving 15 lanes to load the sample on. Protein samples were prepared by mixing 9 µl of protein samples with 3 µl of 4× loading buffer and incubated at 90 °C for 3 min to denature the samples. The gel was placed in a tank filled with 1 × running buffer. 7 µl of samples and PageRuler Plus Prestained Protein Ladder were loaded on the gel. Samples were separated at a constant voltage of 180 V for 1 hour in a 1 × SDS running buffer.

10 × SDS running buffer: Tris base (30 g), glycine (144 g) and SDS (10 g) in 1000 ml of deionised water.

Table 2-8. Ingredients and volumes required for casting SDS-PAGE Gel

Reagents	Volumes (mL) required for 12% resolving gel	Volumes(mL) required for 5% stacking gel
H_2O	1.6	0.68
30% acrylamide mix	2	0.17
1.5 M Tris (pH 8.8)	1.3	—
1.5 M Tris (pH 6.8)	—	0.13
10% SDS	0.05	0.01
10 % ammonium persulfate	0.05	0.01
TEMED	0.002	0.001

2.5.6. Coomassie blue staining of proteins

Coomassie brilliant blue was used for visualization of separated proteins by SDS-PAGE. Following electrophoresis, the gel was washed 3 × with water by microwaving the gel with water at medium power setting of the microwave. This washing step was required to remove the excess of SDS. Then the gel was stained with a solution of Coomassie brilliant blue by

microwaving the gel with the staining solution. The gel was then destined by shaking the gel in water at room temperature overnight.

2.5.7. Buffer exchange and centrifugal concentration of proteins

The fractions containing desired protein were combined and placed into vivaspin 20 centrifugal column (Fisher Scientific). The molecular weight cut off (MWCO) was selected to be 50% smaller than the molecular size of the protein. The column was previously washed with water. The protein sample was washed 3× with the storage buffer. When it was necessary, the protein sample was further concentrated using the same column. The protein samples were aliquoted and stored at -80 °C until they needed.

2.5.8. Quantification of proteins

The concentration of protein was measured using NanoDrop ND-1000 UV spectrophotometer. The absorbance of protein at 280 nm was measured and the concentration of protein was calculated according to Beer-Lambert law using the extinction coefficient of protein as estimated by ExPASy web tool.

2.6. Spectrophotometric determination of enzyme activity: ADH with ethanol as the substrate

In the ADH activity assay using ethanol as the substrate, the rate of reduction of NAD^+ to NADH is monitored spectrophotometrically at 340 nm. One unit of ADH is the amount of enzyme that will generate 1.0 µmole of NADH per minute at pH 8.8 and 25°C.

In this work, commercially available ScADH (from Merck: A7011) was utilised.

Table 2-9.ADH activity set up.

Reagent	Volume (µl) for 200 µl total volume in microplate	Final concentration
Sodium phosphate buffer (50 mM, pH 8.8 at 25˚C)	86.6	21.65 mM
Ethanol 95%(V/V)	6.7	3.1%

β-NAD (15 mM) in MilliQ water	100	7.5 mM
Diluted ADH solution (from 1 mg.mL⁻¹ stock)	6.7	
Enzyme Diluent		
(10 mM Sodium phosphate, pH 7.5 with 0.1 %		
(w/v) Bovine Serum Albumin)		

The assay was carried out using volumes and concentrations as specified in Table 2-9. The background rate was measured for 1 min before initiating the reaction with the addition of the substrate, ethanol. The reaction rate was measured over 2 min.

Specific acitiviy was calculated the linear part of the slope and Equation 2-1:

<p style="text-align:center;color:#4a90d9;">Equation 2-1</p>

$$\frac{Units}{mL}\,enzyme = \frac{\left(\frac{A_{340}}{min}\,reaction\right) - \left(\frac{A_{340}}{min}\,blank\right)(200)(df)}{(\varepsilon)(L)(6.7)}$$

$$\frac{Units}{mg}\,enzyme = \frac{\frac{Units}{mL}\,enzyme}{\frac{mg}{mL}\,enzyme}$$

200=Total volume (in µl) of assay

df = Dilution factor (a factor by which the stock solution of enzyme is diluted)

ε =Extinction coefficient of ß-NADH at 340nm = 6.22 µmol⁻¹ mL cm⁻¹ [49]

6.7 = Volume (in µl) of enzyme used

L=Path length of the microplate for a 200 µL well volume= 0.58 cm

2.7. Spectrophotometric determination of *Sc*PDC enzyme activity

One unit of *Sc*PDC decarboxylase activity is defined as the amount of *Sc*PDC which catalyses the decarboxylation of 1 µmol of pyruvate per minute under the standard condition (pH 6.0, 30 ˚C).

Table 2-10 summerizes the volumes and concentrations used in the assay. The assay was performed in UV transparent 96 well plate. The assay was perfomed for 350 seconds at 25°C. A mastermix of citrate, sodium pyruvate, NADH and *Sc*ADH was prepared before. The volume of 194 μl of master mix was placed in each well. The reaction was started by adding *Sc*PDC.

Table 2-10. Reaction set up for *Sc*PDC decarboxylase activity assay by couple assay, Temperature of assay 30 °C, pH 6.0 and time 350 s

Reagents	Initial concertation	Volume (μl) for 200 μl total volume in microplate	Final concentration
Citric acid	0.2 M, pH 6	181.5	181.3 mM
Sodium Pyruvate	1 M	6.25	31.3 mM
NADH	21 mM	3.125	0.32 mM
*Sc*ADH	1 mg/ml	3.125	
*Sc*PDC	0.23 mg/ml	6	

Specific acitiviy was calculated following the Equation 2-1.

In the direct decarboxylase assay the decompostion of substrate was monitored at 340 nm by the addition of decarboxylase enzyme. The reaction was set up as below

Table 2-11. Direct decarboxylase assay

Reagents	Initial concentration	Volume (μl) for 200 μl total volume in microplate	Final concentration
Substrate in citrate buffer	33 mM	190	31.3 mM
*Ll*KdcA	0.14 mg/ml	10	

2.8. Determination of kinetic parameters for recombinant pyruvate (de)carboxylases

Purified *Sc*PDC was analysed for decarboxylation reaction using alcohol dehydrogenase as described above. The steady state parameters were obtained by varying the concentration of pyruvate from 0.2 to 31.25 mM. The initial rate (V_0) was plotted against substrate concentration [S]. The kinetic data were obtained by fitting the data in (Equation 2-2) using non-linear regression (curve fit), the allosteric sigmoidal option in Prism 7, where the Hill equation was used.

Equation 2-2

$$V = \frac{V_{max}[S]n}{(K_m)^n + [S]^n}$$

Where n=1 in Michaelis Menten equation

$$V = \frac{V_{max}[S]}{(K_m) + [S]}$$

2.9. Site directed Mutagenesis

Oligonucleotide-directed mutagenesis is a commonly used technique *in vitro* in order to deliver the desired mutation at the specific position. Using a Q5® Site-Directed Mutagenesis kit (NEB), a pair of mutagenic primers are designed and synthesised. The mutagenic primer which contains DNA bases complementary to a specific region in a cloned gene except for a mismatch at the target codon which directs the desired mutation. [50]

Table 2-12.Primes used for Q5 Site-Directed Mutagenesis. Codons carrying the mutations are printed bold and italic.

Mutation	5'-Forward primer -3'	5'-Reverse primer-3'
PDC_E504Q	TTACACCATT***cag***AAGTTGATTCACG	CCATCGTTGTTCAAGACG
PDC_I507A	TGAAAAGTTG***gcg***CACGGTCCAAAG	ATGGTGTAACCATCGTTG
PDC_T415S	CATTGCTGAA***agc***GGTACCTCCG	ACAACATCACCTTCTTGCAAGAAG
PDC_T415A	CATTGCTGAA**gcg**GGTACCTCCG	ACAACATCACCTTCTTGC
PDC_F319A	GTTGTCTGAT**gcg**AACACCGGTTCTTTC	AAAGCACCGACAGACAAAATC

Table 2-13.PCR reaction setup

PCR reaction reagents	Volume (µl) for 25 µl reaction volume	Final Con.
Q5 Hot start High fidelity 2X Master Mix	12.5	1x
10 µM Forward Primer	1.25	0.5 µM
10 µM Forward Primer	1.25	0.5 µM
Template DNA	1	15 ng µL^{-1}
Nuclease-free water	9	-

The Q5 High-Fidelity Master Mix contains 2.0 mM Mg^{2+} when used at a 1X concentration. The final concentration of dNTPs is 200 µM of each deoxynucleotide in the 1X Q5 High-Fidelity Master Mix. The concentration of Q5 High-Fidelity DNA Polymerase in the Q5 High-Fidelity 2X Master Mix has been optimized for best results under a wide range of conditions.

Table 2-14. Thermocycling conditions for PCR reaction

Step	Temperature (°C)	Time (s)
Initial Denaturation	98	30
25 cycles	98	10
	Annealing temperature	30
	72	(30 seconds/kb=120)
Final Extension	72	120
Hold	4-10	

2.9.1. Kinase, Ligase and DpnI (KLD) reaction

The KLD mixture contains a mix of kinase, ligase and DpnI enzymes in a buffer that retains the activity of the enzymes. The PCR derived plasmids were treated with Kinase and ligase in order to phosphorylate both ends and circularize them. The template DNA is removed with DpnI enzyme, since it digests methylated and semi-methylated DNA.

Table 2-15.KLD reaction setup

KLD reaction reagents	Volume (µL)	Final Con.
PCR Product	1	
2X KLD reaction buffer	5	1X
10X KLD enzyme mix	1	1X
Nuclease-free water	3	-

The KLD reaction was setup according to Table 2-15. The reaction mixture was mixed by pipetting up and down and incubated at room temperature for 10 minutes followed by additional 30 minutes on ice.

2.10. Sequencing

Plasmid DNA diluted in nuclease free water was submitted to DeepSeq sequencing at University of Nottingham for DNA sequencing and confirmation of mutations.

2.11. One-pot carboxylation-transamination reaction

The reaction was set up inside a 1 ml pressure reactor (PressureSyn, University of Nottingham in collaboration with Asynt) using acetaldehyde or propanal (25 mM), TPP (1 mM), $MgSO_4$ (2 mM), L-glutamate (50 mM), ScPDC (20 μM) and ScALT (30 μM) to a final volume of 1 ml. The reaction was started by applying CO_2 at pressure of 55 psi ($\cong$ 4 bar). To investigate the effect of pH on the reaction, the reaction was performed at various buffers (250 mM) with different pH's: sodium phosphate (NaH_2PO_4-Na_2HPO_4) (pH 7.0_8.0), 250 mM HEPES (pH 7.5) and 250 mM sodium carbonate ($NaHCO_3$-Na_2CO_2) (pH 8.5_10.5). Also, the effect of pressure on the reaction was tested by setting up the reaction using the conditions as described above, using $NaHCO_3$ (250 mM, pH 8.5) and various pressures of CO_2 (70 psi,0.48 MPa), (140 psi, 0.96 MPa) and (200 psi, 1.37 MPa). The reaction mixture was incubated at 25 ℃ and stirred at 100 rpm for 48 hours. The formation of L-alanine and L-homoalanine was analysed by HPLC. In the carboxylation reaction catalysed by ScPDC the reaction conditions were exactly as above without the addition of L-glutamate and ScALT.

2.12. Calculation of the $[CO_2{}^{aq}]/[HCO_3^-]$ ratio at pH 8.5 [29]

$$CO_2{}^{aq} \xrightleftharpoons{H_2O} H_2CO_3 \qquad K_h = \frac{[H_2CO_3]}{[CO_2{}^{aq}]} = 1.7 \times 10^{-3}$$

$$H_2CO_3 \rightleftharpoons HCO_3^- + H^+ \qquad K_a = \frac{[HCO_3^-][H^+]}{[H_2CO_3]} = 2.5 \times 10^{-4}\,M$$
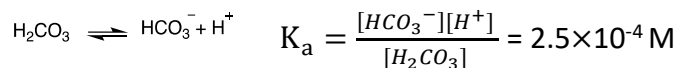
From above equations

$$[H_2CO_3] = [CO_2{}^{aq}] \times (1.7 \times 10^{-3}) \qquad pH = -\log[H^+]$$

$$\frac{[HCO_3^-][10^{-8.5}]}{[CO_2{}^{aq}] \times (1.7 \times 10^{-3})} = (2.5 \times 10^{-4})$$

$$\frac{[CO_2{}^{aq}]}{[HCO_3^-]} = 7.2 \times 10^{-3}$$

To calculate the $[CO_2{}^{aq}]$ at the pressure of 70 psi

According Henry's law,[29]

$$CO_2{}^{gas} \xrightleftharpoons{K_H} CO_2{}^{aq} \qquad \frac{1}{[K_H]} = \frac{[CO_2{}^{aq}]}{p_{CO_2}} = 33.6\ mM.bar^{-1}$$
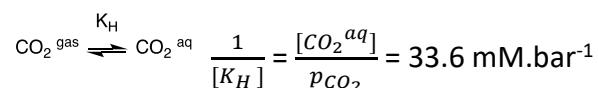
Therefore, at pressure of 70 psi (4.8 bar) the concentration of CO_2 is 162 mM.

2.13. HPLC detection of α-keto acid

Keto acids (pyruvate and 2-ketobutryic acid) were analysed by the use of a Dionex Ultimate 3000 HPLC system equipped with Rezex ROA-Organic Acid H⁺ (8%) 150 mm×7.8 mm×8 μm column, a refractive index (RI) and diode array detector (DAD) at UV 210 nm. The mobile phase consisted of 0.005 M H_2SO_4 and an isocratic flow rate of 0.5 ml/min for 30 minutes. The column temperature was 35 °C. Samples and standards were prepared by mixing with valeric acid as an internal standard (80 mM in 0.005 M H_2SO_4) in a 9:1 ratio prior to the injection (20μl) into the column.

2.14. HPLC detection of amino acids: Aminotransferase HPLC activity assay

The aminotransferase activity of alanine transaminase from *Saccharomyces cerevisiae* (*Sc*ALT) was measured by HPLC. This was performed by measuring the amount of the expected amino product and calculating the percentage of the activity.[51] An aliquot of the enzyme (30 μM) was incubated with each 2-keto acid (pyruvate or 2-ketobutryic acid, 2.5 mM) and glutamate (25 mM) as the amino donor and acceptor, respectively. The reactions were carried out in 1 mL at 25°C in $NaHCO_3$-Na_2CO_3buffer (250 mM, pH 8.5) with shaking at 100 RPM. After the incubation period (24,48 and 72 hours), an aliquot of the reaction mixture was removed and centrifuged at 13000 rpm for to remove any precipitate. Pre-column derivatisation was carried out prior to analysis by HPLC.

2.14.1. Precolumn derivatisation step with 9-fluorenylmethyl chlorocarbonate (FMOC-Cl)

The derivatisation step was conducted by mixing 25 μl of the reaction mixture with 100 μl of FMOC-Cl (15 mM in Acetonitrile) and 50 μl of borate buffer (100 mM, pH 9). The mixture was left at room temperature for 5 min and centrifuged at 13000 rpm for 1 minute to remove any precipitate formed. To 175 μl of the cleared supernatant, 400 μl of acetonitrile and 400 μl of 0.2% HCl were added. Amino acid formation was analysed by injecting 2 μl of above mixture into a XBridge C18 column (2.1 mm× 100 mm×3.5 μm column) followed by a gradient elution and detection at 280nm. The elution program was as follow: 5-95% acetonitrile supplemented with 0.1 % trifluoracetic acid for 4 min at the flow rate of 1 ml/min.

The yields were calculated using calibration plots. The calibration plots were obtained by derivatising and injecting the standard solutions of each amino acids (alanine or homoalanine) with varying concentrations (0, 0.5, 1 and 2 mM).

2.14.2. Precolumn derivatisation step with *O*-Phthaldialdehyde (OPA)

The procedure was adapted from Hanko *et al*.[52] OPA solution (6 mM, 1ml) was prepared in HPLC-grade methanol. The above solution was diluted in 100 ml solution of boric acid (0.4 M) (pH to 10.4 by 5M potassium hydroxide) supplemented with β-mercaptoethanol (200 μl). The samples (either standards or reaction solution) were prepared by adding 950 μl of 50%

methanol to 50 μl of each sample. In vial derivatisation was performed by mixing 50 μl of amino acid samples with 50 µl of OPA solution. The concentration of amino acids was analysed by injecting 5 µl of above mixture into a Kintex EVO C18 followed by a gradient elution and detection at 338nm. Two mobile phases were used: A) 20 mM KH_2PO_4 adjusted to pH 7.2 with KOH and B) methanol/acetonitrile (50:50 v/v). A gradient elution program was followed as below with a flow rate 1 ml/min for 23 min.

Table 2-16. Gradient elution program

Time (min)	%A	%B
0	97	3
20	40	60
23	97	3

2.15. Screening the selectivity of commercially available ω-transaminases against different aldehydes and their corresponding 2-keto acids

Biotranfmormations of different aldehydes (acetaldehyde, propanal, benzaldehyde and phenyl aldehyde) and their corresponding 2-keto acids (pyruvate, 2-ketobutryic acid, phenylglyoxylic acid and phenylpyruvic acid) were carried out using different ω-TAs (ATA-113, ATA-025, ATA-0117, ATA-256). A solution containing PLP (1 mM) and xylylenediamine dihydrochloride (Sigma-Aldrich, product code 711462) (46.4 mg, 11.1 mM) in the carbonate buffer (20 mL, 100 mM, pH 10.0) was prepared. Each ω-TA (0.5 mg/mL) was rehydrated in 5 ml of above solution. Enzyme containing solution (180 μl) was dispensed into a well of a 96-well plate. Aldehydes or their corresponding 2-keto acids (20 mM, 20 μl from a 200 mM stock in DMSO) were added separately to each well. The plate was incubated at 30℃ with shaking 150 rpm for 24h. In a control experiment DMSO was added instead of substrate (aldehyde or 2-keto acid). Colour changes were observed.

2.16. Chemical synthesis of propanal bisulfite

In a 50 ml round bottom flask a solution of propanal (10 mmol in 20 mL of methanol) was prepared. A solution of sodium bisulfite (10 mmol in 3 ml water) was gradually added to the propanal solution. The mixture was stirred at room temperature for 3 days. The solvent was removed under the pressure using a rotary evaporator. The water removal was further assisted by adding 5 ml of methanol and repeating the evaporation. A white solid (1.49 g, 92%) was collected and washed with diethyl ether (8 ml). ^{1}HNMR (D_2O) = δ 4.20-4.24 (dd, 1H), 1.87-1.94 (dq, 1H), 1.42-1.52 (m, 1H), 0.95-0.99 (t, 3H) ppm. The mass spectrum of the propanal bisulfite in H_2O was recorded on a Bruker Micro TOF mass spectrometer showing the $Na_2[C_3H_8NSO_3]$ with size of 184.98 Da.

Chapter 3. Preparation of a thiamine dependent decarboxylase enzyme pyruvate decarboxylase from *Saccharomyces cerevisiae*

This chapter provides a summary covering pyruvate decarboxylase (PDC), its mechanism and structure. In the experimental part, PDC-pJ414 expression vector is constructed by amplifying the PDC gene from *S. cerevisiae* and insertion into pJ414 expression vector. Finally, PDC is expressed and purified to investigate its reverse reaction for the fixation of CO_2.

3.1. Background

PDC (EC 4.1.1.1) is the enzyme involved in alcoholic fermentation which catalyses the conversion of pyruvate to acetaldehyde and CO_2 (Scheme 3-1, A). In 1911, the decarboxylase activity of PDC by fermenting yeast was described by Neuberg *et al.*[53] Efficient enzymatic catalysis by PDC requires both Thiamine diphosphate (ThDP) and magnesium ions (Mg^{2+}).[54]

PDC can catalyse two different reactions: A) non-oxidative decarboxylation of α-ketoacids and B) carboligase side-reaction resulting in the formation of α-hydroxyketones (Scheme 3-1, B). For many years, PDCs from various organisms have been studied mainly with respect to the mechanism of the non-oxidative decarboxylation. Though the C-C bond forming characters of these enzymes are known and have been used in biotransformations for many years for the synthesis of α-hydroxyketones, only a little is known about the factors affecting the carboligase side reaction.[55]
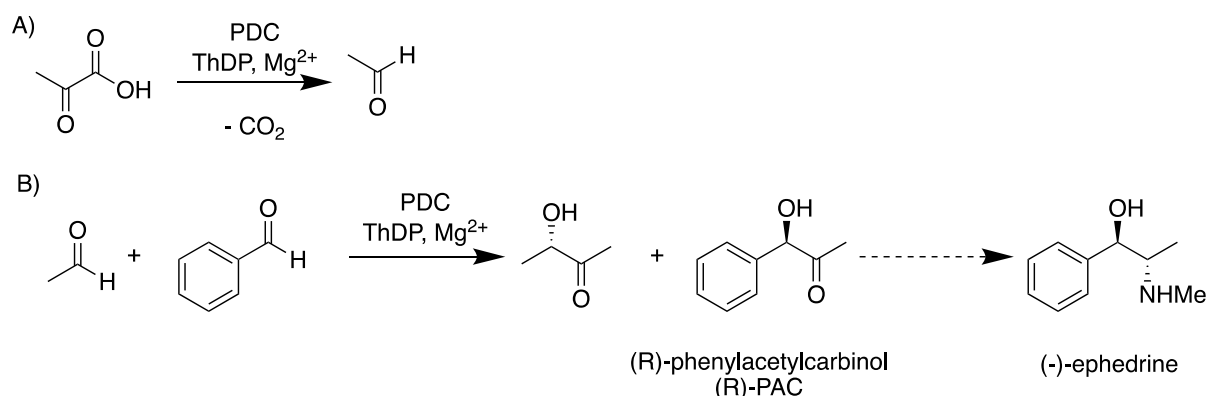
Scheme 3-1. Reactions of PDC A) non-oxidative decarboxylation B) carboligation reaction.

3.1.1. Thiamine diphosphate (ThDP)

Thiamine diphosphate (ThDP), the biologically active form of vitamin B_1, is a common cofactor for many different enzymes involved in metabolic pathways. It consists of a pyrimidine ring, a thiazolium ring and a methylene group that bridges the two rings together (Scheme 3-2, B). Generally, ThDP-dependent enzymes are involved in the formation (ligase) and cleavage (lyase) of carbon-carbon bonds adjacent to a carbonyl group.[56]
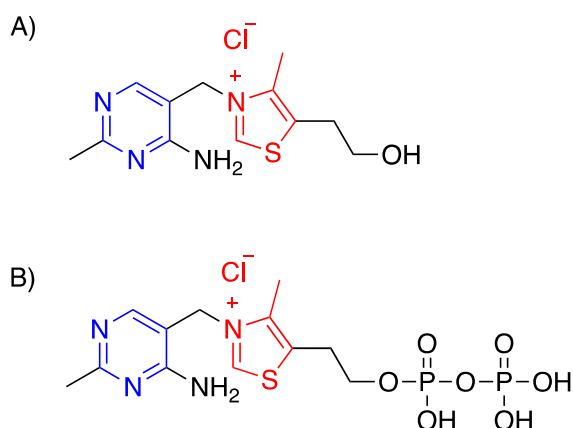
Scheme 3-2. Schematic drawing of A) vitamin B_1 and B) thiamine diphosphate

3.1.2. PDC: Mechanism of non-oxidative decarboxylation of α-ketoacids

PDC catalyses the non-oxidative decarboxylation of α-ketoacids. In this mechanism, the first step in the catalytic process is formation of the ThDP thiazolium ylide.
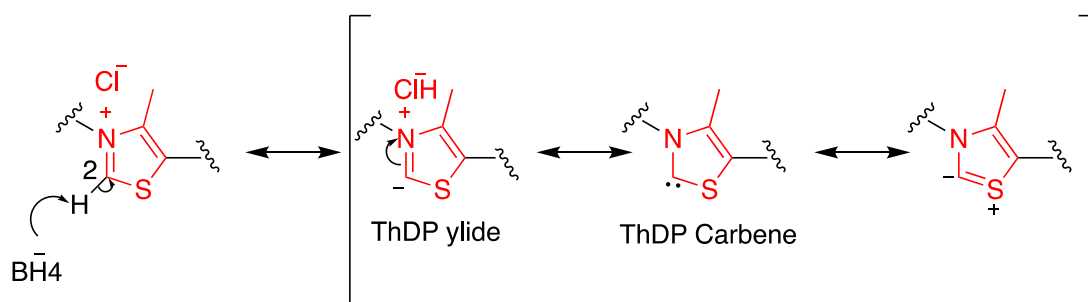
Scheme 3-3. Resonance stabilisation of thiazolium ylide

The acidity of the C2 proton (pK_a 12.1) [57] on the thiazolium ring has been attributed to an electrostatic interaction (ylide formation), delocalization (carbene formation), and *d-p* orbital overlap with the sulphur atom (Scheme 3-3).[58]

In the decarboxylation mechanism, the cofactor is held in a ''V'' conformation which brings the hydrogen atom at the C2 position close to the nitrogen atom at N4', approximately within 3-3.4 A° (Scheme 3-4).[59]
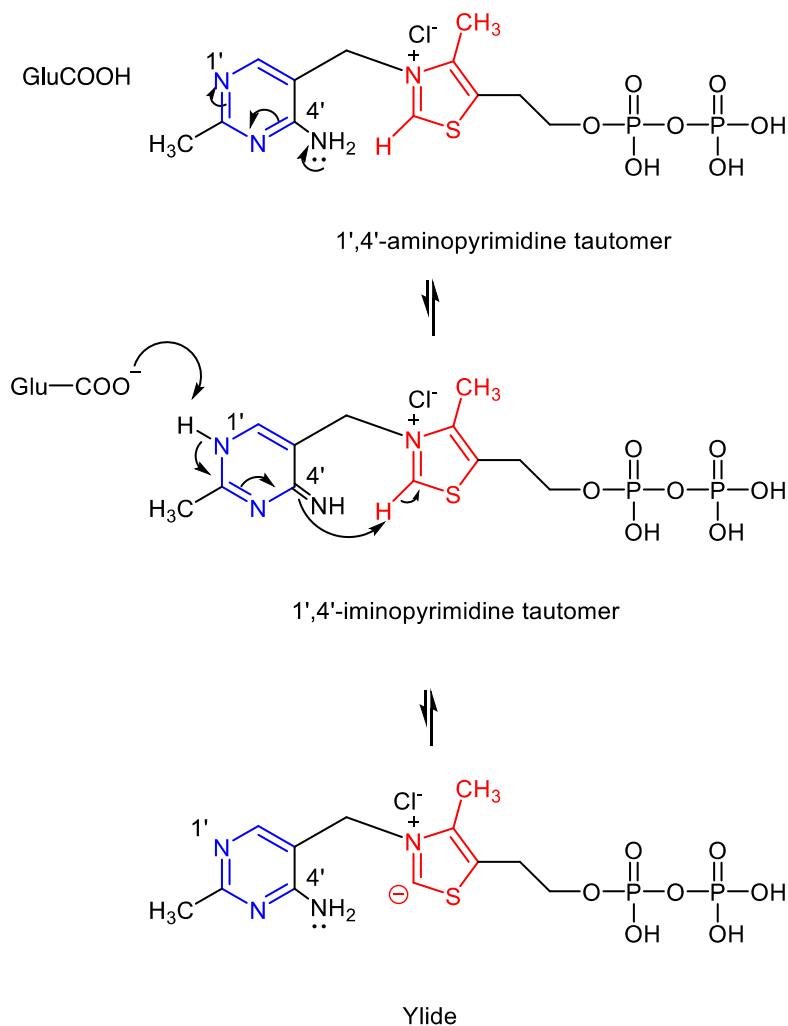
Scheme 3-4. PDC decarboxylation mechanism, formation of thiazolium ylide of ThDP

The first step in the reaction mechanism is the nucleophilic attack of the ThDP ylide at the α-carbonyl group of the keto acid substrate resulting in the formation of the 2α-lactyl-ThDP. This step is followed by the loss of CO_2 and the formation of an enamine intermediate, which sometimes called ''activated aldehyde intermediate''. Next, the enamine becomes protonated to form the 2α-hydroxyethyl-ThDP (Breslow) intermediate. Finally, the product of the decarboxylation, the aldehyde, is released to regenerate the ThDP (Scheme 3-5).
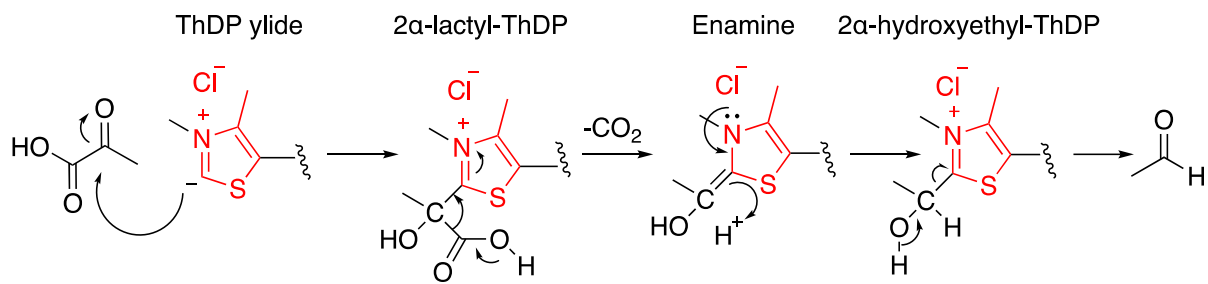
Scheme 3-5. Mechanism of decarboxylation of pyruvate to acetaldehyde catalysed by PDC

3.1.3. 3D-structure of PDCs from different organisms

The three dimensional structure of the α_4 isoenzyme of PDC *from S. uvarum* (*Su*PDC) was reported for the first time in 1993.[60] Three years later, the 3D structure of the isoenzyme α_4 of PDC from *S. cerevisiae* (*Sc*PDC) was published (Figure 3-1).[44] All the PDCs from yeast and bacteria are tetramers in the native state and consist of identical or almost identical subunits with molecular masses of approximately 60 kDa.
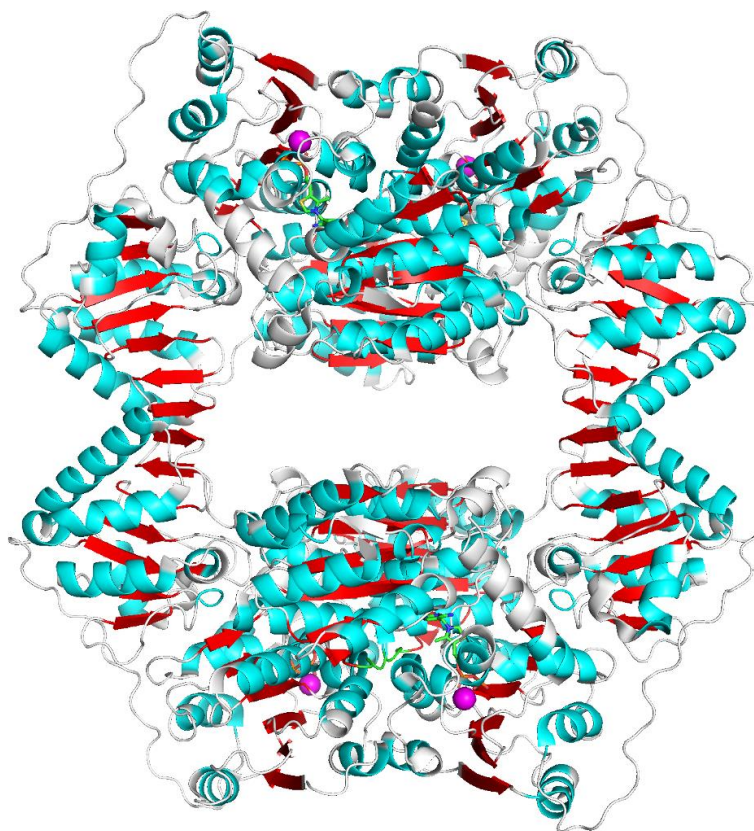
Figure 3-1. Structure of *Sc*PDC (PDB accession number: 1PVD)[44] as a tetramer generated by PyMOL

Plant PDCs are similar to those from yeast and bacteria but they are organised in higher oligomers.[61] Molecular masses of 300-500 kDa have been reported for PDC from rice, pea and wheat.[62-64] Each subunit of PDC contains a single polypeptide chain which is organised in three structurally compact domains: 1) α-domain, the amino-terminal domain that non-covalently binds to the pyrimidine ring of ThDP 2) γ-domain, the carboxy-terminal domain which forms the diphosphate binding site 3) β-domain, the middle domain that provides the intermolecular contact between two dimeric halves in the tetramer and has the substrate regulation site in PDCs.

3.1.4. 3D-Structure of the *Sc*PDC subunit

In *Sc*PDC, the α-domain (residues 1 to 187) and γ-domain (residues 360 to 556) have similar topology with a central six-strand parallel *beta* sheet surrounded by six α helices. Similar topology also exists in the other ThDP-dependent families including transketolase (TK), oxoreductase (OX) and 2-ketoacid dehydrogenase (KD).[65, 66] The β-domain has different topology compared to the two other domains. It contains seven-strand mixed sheets of five strands which are parallel and the other two are antiparallel (Figure 3-2).

α −domain   γ −domain   β −domain
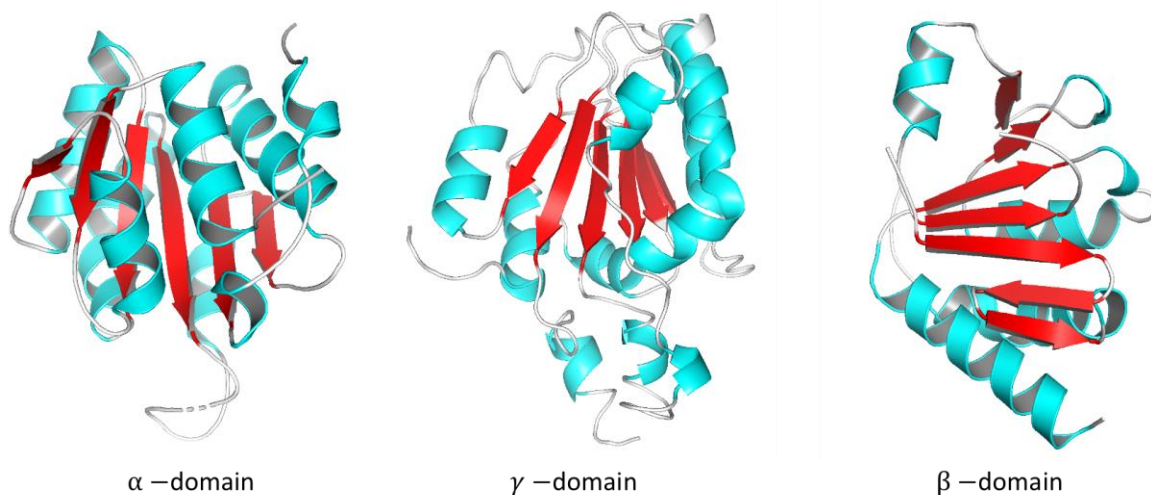
Figure 3-2. Comparison between the α, γ and β domains topology in *Sc*PDC subunit, generated by PyMOL

The ThDP cofactor is, for the most part, bound to the γ-domain, but also has contacts with some residues (P26 and, most importantly, E51) in the α-domain of a neighbouring subunit (Figure 3-3).
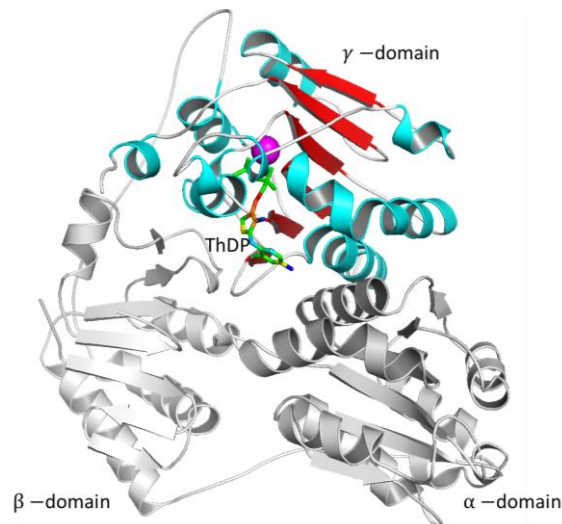
Figure 3-3. The ThDP in a single subunit of ScPDC, generated by PyMOL.

The active sites, where the cofactors ThDP and Mg^{2+} are present, are located between the α-domain of one subunit and the γ-domain of the other subunit within a dimer ($α_1γ_2$ and $α_2γ_1$). There are two active site per dimer and four per tetramer (Figure 3-4). The cofactor binding cavity is roughly 10 A° wide and 8 A° deep.
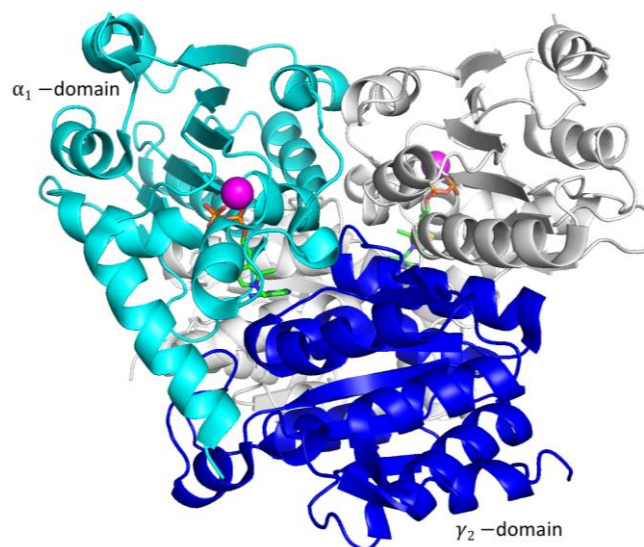
Figure 3-4. Representation of the active site in PDC dimer, generated by PyMOL[67]

The Mg^{2+} ion provides a platform for binding of the ThDP. The cofactor is stabilised through the formation of an inner sphere complex. This octahedral complex is formed by the coordination of the two oxygen atoms from the phosphate group of ThDP and three other amino acids including: D444-OD1, N471-OD1, G473-O with Mg^{2+} ion (Figure 3-5).
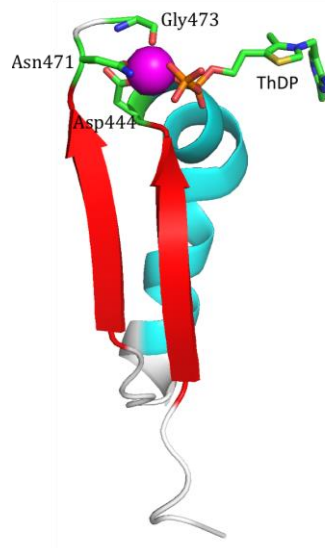
Figure 3-5. The octahedral complex of Mg^{2+} ion in the active site

There is a common sequence in ThDP-dependent enzymes, including pyruvate decarboxylase from *S. cerevisiae*. The common sequence motif starts with the highly conserved sequence, GDG triplet, and terminates with the highly conserved sequence, NN doublet (Figure 3-6).[68] In between there are roughly 30 residues and their sequence is much less conserved. These residues are among the few highly conserved in all ThDP-dependent enzymes and form the Mg^{2+} binding region that helps in anchoring the diphosphate of the ThDP to the protein. This region around the diphosphate-Mg^{2+} binding site is called ''fingerprint''.

443   473

GDGSLQLTVQEISTMIRWGLKPYLFVL**NN**DG

Figure 3-6. The fingerprint region in *Sc*PDC

3.2. Results and discussion

3.2.1. Cloning of *Sc*PDC into the pJexpress414 vector

Cloning of the *Sc*PDC gene required a source of genomic DNA. Isolation of the genomic DNA of *S. cerevisiae* was obtained by the glass-bead method and used throughout this work. This protocol included three steps: Lysis, precipitation and purification. The cells were mechanically disrupted in the lysis buffer using glass-beads. The precipitation of the DNA was achieved by addition of ammonium acetate (7 M solution, pH 7) and isopropanol. To purify the DNA, the pellets were washed with ethanol (70%) to remove any unwanted materials and cellular debris. Finally, the pellets were resuspended in sterile water and stored at -80 ˚C. The concentration of *S. cerevisiae* genomic DNA was determined by measuring UV absorbance at 260 nm and 280 nm. The A_{260}/A_{280} ratio of the DNA, was 1.8, indicating high level of purity of the isolated DNA.

Cloning of the PDC from the *S. cerevisiae* genomic DNA was accomplished using Q5 High-Fidelity 2× Master Mix which contains Q5 High- Fidelity Polymerase, following the manufacture's protocol. The primers (*Sc*PDC-FWD and *Sc*PDC-REV) specifically were designed to amplify the PDC and incorporate restriction sites (*BamHI* and *XhoI*) at the 5' to 3' ends of the gene to allow for subsequent digestion and ligation into pJexpress414 expression vector. The primer concentrations and annealing temperatures (T_a) were varied to see if there is any difference in the PCR amplification. After 35 cycles of amplification the PCR product was examined by 1.2% (w/v) agarose gel (Figure 3-7). All conditions showed a single band of PCR product between 1.5 and 2 kilo base pairs in size (*Sc*PDC:1.7 kb).
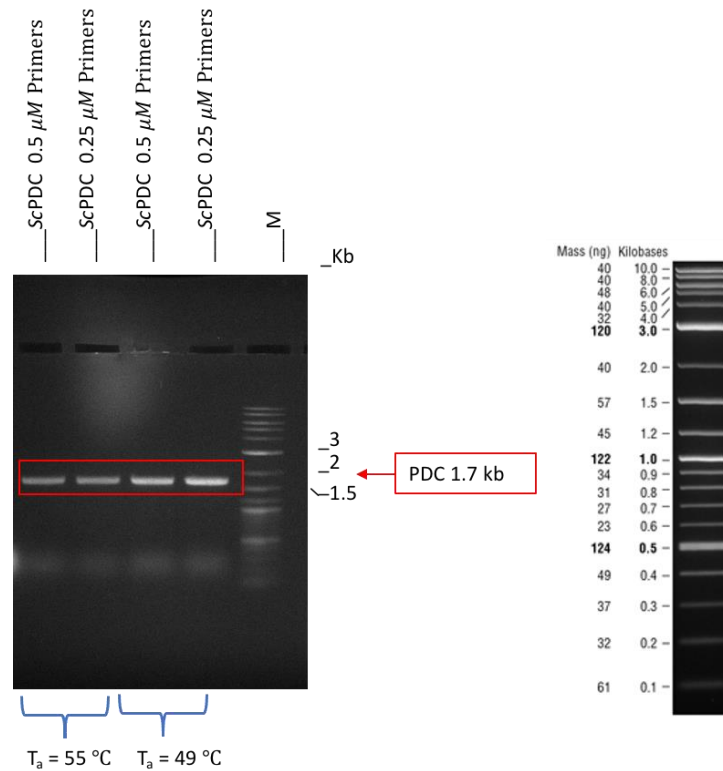
Figure 3-7. The 1.7 kb *Sc*PDC gene fragment amplified by PCR using variable annealing temperatures (55 and 49 °C) and primers concentrations (0.5 and 0.25 µM), M represent the DNA ladder (1kb Plus DNA ladder, N0550G).

The PCR product was purified directly from the reaction mixture using Promega Wizard SV PCR clean up kit according to the manufacturer's protocol to remove the excess of nucleotides and primers. Following the genomic DNA PCR reaction, the purified PDC gene fragment, as well as the pJexpress414 expression vector (with HC ferritin insert given by colleagues at NRT group) were digested by *BamH*I-HF and *XhoI* restriction enzymes at 37°C for 1 hour in preparation for ligation step. The resulting fragments were separated on 1.2% (w/v) agarose gel and visualised under UV light (Figure 3-8). The desired pJexpress414backbone was excised from the gel using a scalpel and the DNA extracted using QIAquick gel extraction kit following the manufacture's protocol.
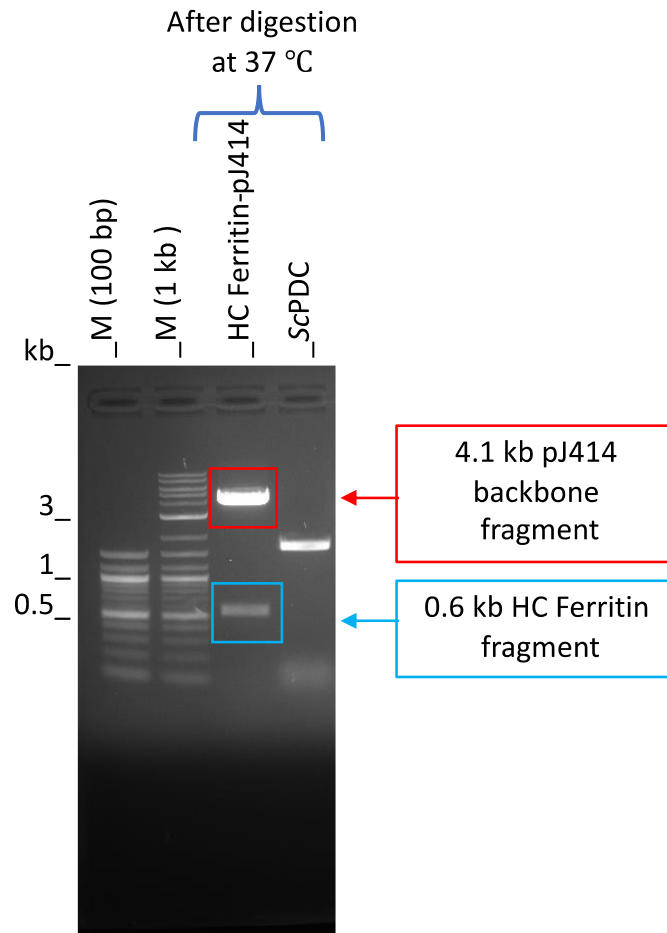
Figure 3-8. Restriction digestion. The digestion of both HC Ferritin-pJexpress414 and *Sc*PDC PCR product was examined by 1.2% agarose gel. M represent the DNA ladder (1kb Plus DNA Ladder, N0550G) and (100 bp DNA Ladder, N3231S).

Both *Sc*PDC fragment and pJexpress414 vector were ligated by T4 DNA Ligase to create circular DNA suitable for transformation into the *E. coli* storage strain (NEB5alpha). The insertion of *Sc*PDC gene into the expression pJ414 vector was confirmed by isolating the DNA plasmid and submitting the sample with T7 forward and pJexpress414Rev primers for sequencing. The DNA sequencing results confirmed that the *Sc*PDC ligation was successful, and the HC ferritin gene was replaced by a 1.7 kb *Sc*PDC gene. The original T7 promoter, His-tag and AviTag™ sequence located at the N-terminal remained in the ligated plasmid (Figure 3-9).

The two tags were designed into the sequence in order to facilitate protein purification. His Tagged proteins can be simply purified by the Immobilised Metal Ion Affinity Chromatography (IMAC) purification method. The Lysine residue within the AviTag can be used to covalently attach biotin to the protein. The biotinylated proteins can be also purified using avidin or

streptavidin.[69] A thrombin cleavage site was incorporated into the sequence in order to remove the His-tag when it is required. The properties of *Sc*PDC were calculated from its amino acid sequence using Expasy tool under ProtParam subsection (Table 3-1).

```
          10         20         30         40         50         60
   MHHHHHHGLN DIFEAQKIEW HELVPRGSSE ITLGKYLFER LKQVNVNTVF GLPGDFNLSL

          70         80         90        100        110        120
   LDKIYEVEGM RWAGNANELN AAYAADGYAR IKGMSCIITT FGVGELSALN GIAGSYAEHV

         130        140        150        160        170        180
   GVLHVVGVPS ISAQAKQLLL HHTLGNGDFT VFHRMSANIS ETTAMITDIA TAPAEIDRCI

         190        200        210        220        230        240
   RTTYVTQRPV YLGLPANLVD LNVPAKLLQT PIDMSLKPND AESEKEVIDT ILALVKDAKN

         250        260        270        280        290        300
   PVILADACCS RHDVKAETKK LIDLTQFPAF VTPMGKGSID EQHPRYGGVY VGTLSKPEVK

         310        320        330        340        350        360
   EAVESADLIL SVGALLSDFN TGSFSYSYKT KNIVEFHSDH MKIRNATFPG VQMKFVLQKL

         370        380        390        400        410        420
   LTTIADAAKG YKPVAVPART PANAAVPAST PLKQEWMWNQ LGNFLQEGDV VIAETGTSAF

         430        440        450        460        470        480
   GINQTTFPNN TYGISQVLWG SIGFTTGATL GAAFAAEEID PKKRVILFIG DGSLQLTVQE

         490        500        510        520        530        540
   ISTMIRWGLK PYLFVLNNDG YTIEKLIHGP KAQYNEIQGW DHLSLLPTFG AKDYETHRVA

         550        560        570        580        590
   TTGEWDKLTQ DKSFNDNSKI RMIEIMLPVF DAPQNLVEQA KLTAATNAKQ
```

Figure 3-9 . Gene design and primary sequence of recombinant *Sc*PDC: His-Tag in red, Avitag in blue, thrombin cleavage site in green and *Sc*PDC in black.

Table 3-1.General properties of a single subunit of *Sc*PDC. The data generated using ProtParam software (Expasy Protemics tools).

	Length (aa)	Molecular weight (kDa)	Theoretical pI	Extinction coefficients at 280 nm ($M^{-1}cm^{-1}$)
ScPDC	590	64.74	5.96	69330

3.2.2. Recombinant over-expression of *Sc*PDC in *E. coli*

The *Sc*PDC-pJexpress414 plasmid was transformed into BL21(DE3) expression strain. Transformation was followed using chemically competent *E. coli* BL21(DE3) cells, plating the resulting culture on LB agar plate containing ampicillin for selective growth. Individual

colonies were picked from LB agar plates and used to inoculate LB medium containing ampicillin. The next day, the starter culture was diluted with fresh LB medium (1/50 to a final volume of 50 mL) and incubated at 37 °C in a shaker incubator until the OD_{595} reached 0.6, at which point, cells were split into two smaller flasks. One of them was induced with IPTG (100 µM final concentration) and the other not. In order to find the optimum temperature and incubation time for the expression of *Sc*PDC three different conditions were examined for the expression of the protein. As can be seen from Figure 3-10, the expression of the soluble protein is maximum when the cells were grown at 20 °C overnight after the induction point. Therefore, this condition was selected for expression and purification for scale-up (1L growth). The optimum expression temperature found in here was also exploited by Martin *et al.* for expression of *Sc*PDC in *E. coli*.[29]
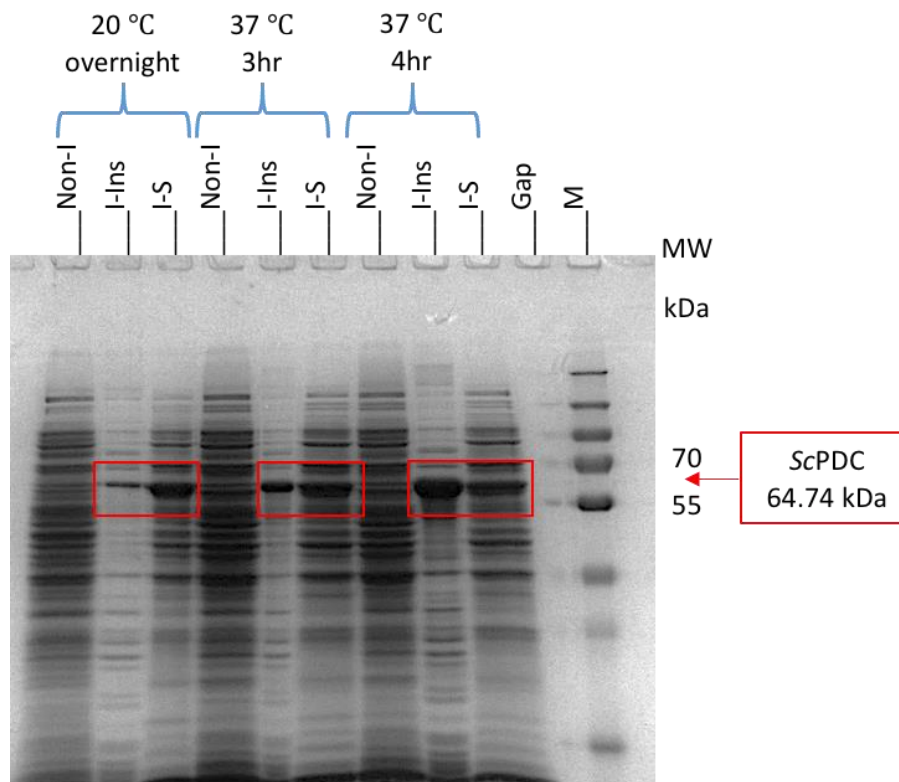
Figure 3-10. SDS-PAGE (12%) analysis of recombinant expression of *Sc*PDC in *E. coli* cells at different conditions. Non-I (Non-Induced), I-Ins (Induced-insoluble), I-S (induced-soluble) and M (Marker: PageRuler™ Plus Prestained Protein Ladder from ThermoFisher).

3.2.3. Purification of recombinant *Sc*PDC

After the optimisation of the condition for expression of *Sc*PDC, the expression was scaled-up to a 1 litre culture. The *E. coli* (BL21) cells were transformed with *Sc*PDC-pJexpress414 plasmid. Recombinant proteins were expressed in LB medium following the addition of IPTG (100 µM) and the growth of cells at 20 ℃ overnight. Cell pellets containing *Sc*PDC protein were centrifuged and stored in $-80\ °C$ until they needed for the purification. The *Sc*PDC was tagged with a N-terminal hexa-histidine-tag, (His)$_6$-*Sc*PDC, giving it a predicted monomer size of 64.74 kDa. For purification, bacterial cell pellets containing (His)$_6$-*Sc*PDC were re-suspended to homogeneity in buffer A and disrupted by sonication. The cell free extract was applied to a Ni-NTA column and the column was washed successively with 5%, 10% and 30% buffer B. His-tagged *Sc*PDC bound proteins were eluted by 50% buffer B. All fractions were analysed by 12% (w/v) SDS-PAGE (Figure 3-11).
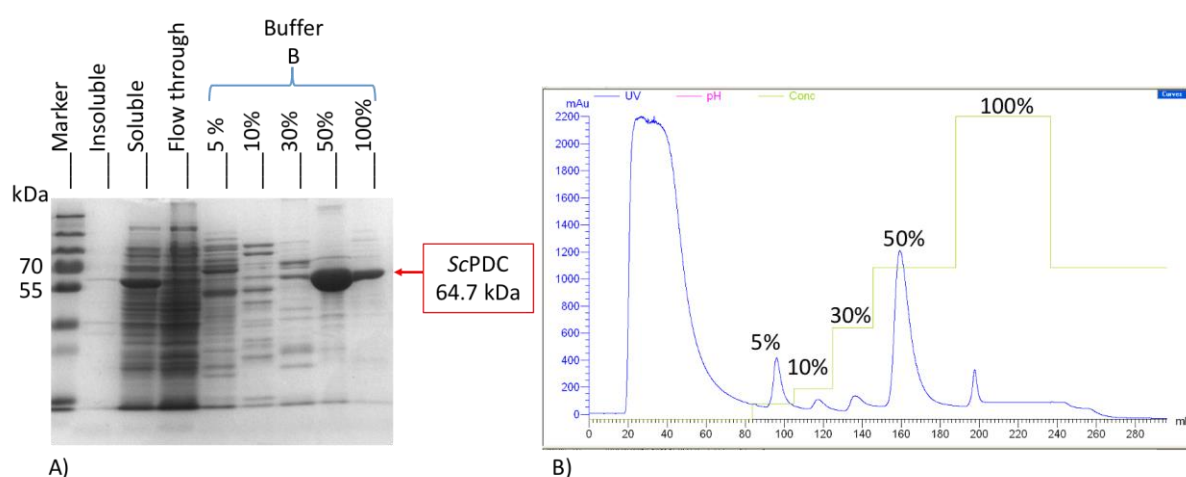
Figure 3-11. Expression and purification of (His)$_6$-*Sc*PDC protein using Ni^{2+} affinity chromatography and AKTA technology. (A) Comassie stained SDS-PAGE (12%) of the proteins present in different fractions including insoluble, soluble, flow through, wash steps and finally elution of *Sc*PDC with different concentrations of imidazole and Marker (PageRuler™ Prestained Protein Ladder) (B) Representative chromatogram for purification of (His)$_6$-*Sc*PDC recombinant protein using Ni^{2+} affinity chromatography and AKTA technology. Absorbance at 280 nm is shown in blue and the concentration of the buffer B used in green.

The protein was buffer exchanged to the dialysis buffer C to remove imidazole at 4˚C overnight. The enzyme solution was concentrated using vivaspin 20 centrifuge column (MWCO 100 kDa). The purity of the stored protein was assessed by SDS-PAGE (Figure 3-12).
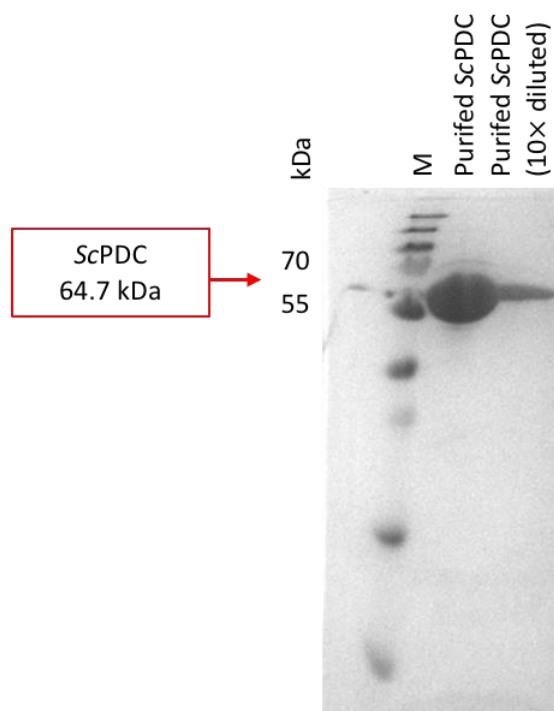
Figure 3-12. SDS-PAGE (12%) of Purified *Sc*PDC, M stands for marker (PageRuler Prestained Protein Ladder).

3.2.4. *Sc*PDC decarboxylase activity by the coupled decarboxylase assay

The decarboxylase activity of the *Sc*PDC can be measured by a coupled assay (Scheme 3-7). In this assay, the product of decarboxylation reaction catalysed by *Sc*PDC is reduced to an alcohol by a second enzyme, alcohol dehydrogenase (ADH). In order to determine the decarboxylase activity of *Sc*PDC, ADH from the same organism (*Sc*ADH) is used as the coupling enzyme.[70] In this work, commercially available *Sc*ADH (from Merck: A7011) was applied in the coupled assay. Therefore, it was important to test the activity of the *Sc*ADH coupling enzyme prior to the coupled assay measurements.

In the *Sc*ADH activity assay using ethanol as the substrate, the rate of reduction of NAD^+ (nicotinamide adenine dinucleotide oxidised form) to NADH was monitored spectrophotometrically at 340 nm (Scheme 3-6). One unit of ADH is the amount of enzyme that will generate 1.0 μmole of NADH per minute at pH 8.8 at 25 °C.[71]
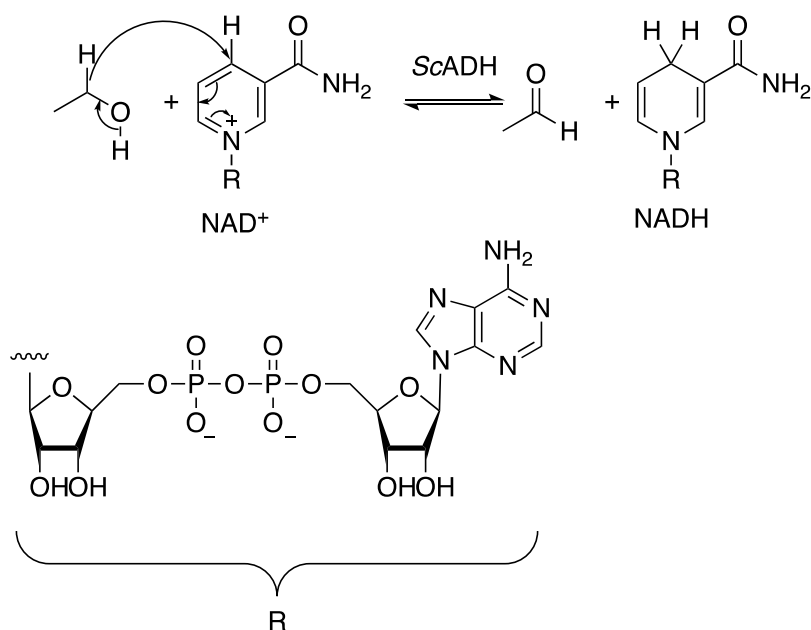
Scheme 3-6. The activity assay of *Sc*ADH followed at 340 nm by the reduction of NAD$^+$ to NADH

Specific activity (μmol/min/mg) of *Sc*ADH was measured following the experimental procedure in section 2.6 and calculated using Equation 2-1.

The pathlength for sample volumes ranging from 25 to 250 μl, based on plate type, have been calculated and reported for common microplates. As pathlength values are proportional to the volume of the liquid used, a linear regression has been reported and can be used to calculate the pathlength of any volume where x=volume used and Y= pathlength (Table 3-2).[72]

Table 3-2. Calculated pathlength for 200 μL volume in a Corning 96 well UV plate

UV compatible plates	Part number	X= Volume (μL)	Linear Regression	Calculated Pathlength [72] cm
Corning 96 well UV plate	3635	200	Y=0.0029x +3e^{-16}	0.58

A serial dilution of the enzyme from stock solution (1 mg.mL^{-1}) was performed to obtain a linear relationship between absorbance at 340 nm and time. At 0.008 mg/ml concentration of *Sc*ADH a linear relationship was established (Figure 3-13). The specific activity was calculated following Equation 2-1.
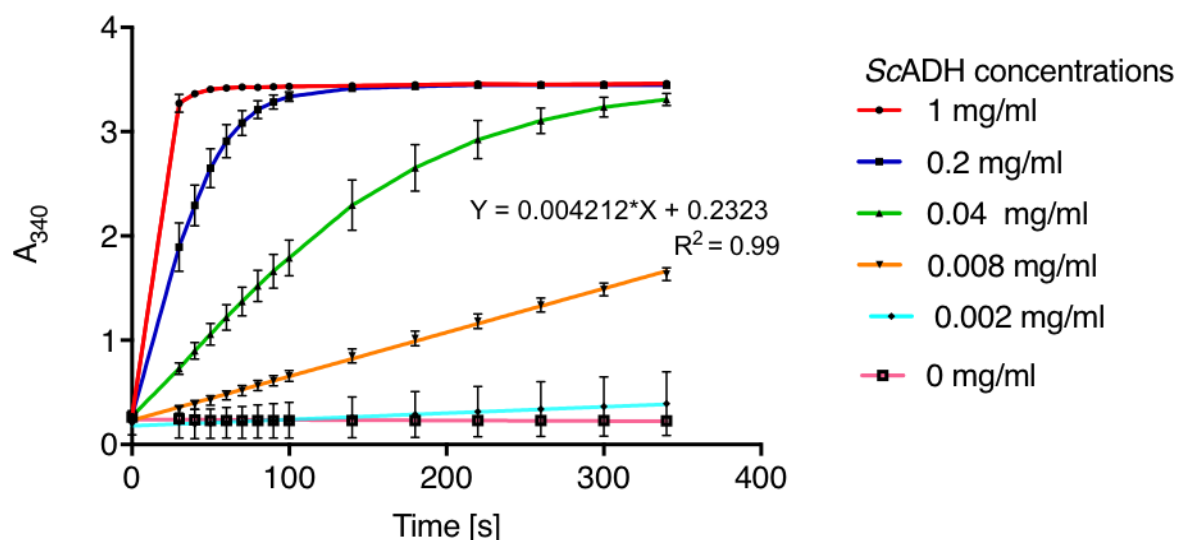
Figure 3-13. Determination of activity of *Sc*ADH.

$$\frac{Units}{mL}\ enzyme = \frac{(0.004127\ S^{-1})(\frac{60\ s}{1\ min})(200)(125)}{(6.22\ \mu mol^{-1}\ mL\ cm^{-1})(0.58\ cm)(6.7)} = 256.1$$

$$\frac{Units}{mg}\ enzyme = \frac{256.1(\frac{Units}{mL}enzyme)}{1\ (\frac{mg}{mL}enzyme)} = 256.1\ \frac{Units}{mg}$$

The obtained specific activity for commercial *Sc*ADH in here was slightly lower compared to the reported value by the manufacture ($\geq$300 units/mg).

Next, the *Sc*PDC activity was measured using the coupled assay as described in section (2.7). One unit of PDC is defined as the amount of enzyme which catalyses the decarboxylation of 1 μmol of pyruvate per at pH 6.0 and 30 ˚C.[73]

In the coupled decarboxylase assay, the aldehyde obtained from the decarboxylation of pyruvate by *Sc*PDC, subsequently is reduced to ethanol by *Sc*ADH. In the coupled assay, the simultaneous rate of NADH oxidation to NAD$^+$ is directly linked to the PDC activity where the reduction step is faster compared to the decarboxylation reaction. At 1mg.mL^{-1} *Sc*ADH concentration the reaction is completed in less than 100 S (Figure 3-13).

Accordingly, one mole of NADH is oxidised to NAD$^+$ in the presence of one mole of acetaldehyde produced (Scheme 3-7). The rate of NADH oxidation is monitored spectrophotometrically at 340 nm.
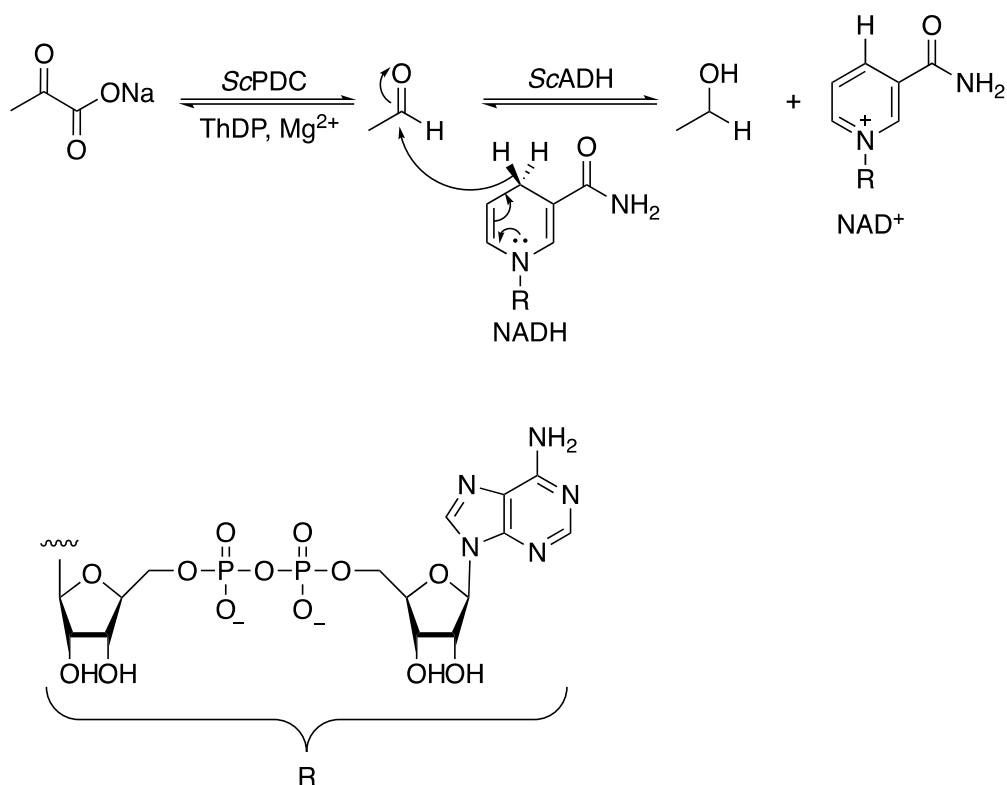
Scheme 3-7. Determination of the decarboxylase activity of *Sc*PDC by the coupled assay

The reaction was initiated by addition of the purified (His)$_6$-*Sc*PDC (6 μL) at five different concentrations.The rate of consumption of NADH was monitored at 340 nm for 350 s. The linear decay of NADH was observed at 0.028 mg/mL concentration of (His)$_6$-*Sc*PDC (Figure 3-14). The rate of NADH decay was calculated by linear regression. The slope of the plot was taken for the specific activity calculationsEquation 2-1. The specific activity of 44.22 units/mg was obtained for (His)$_6$-*Sc*PDC which was in good agreement with the reported value elsewhere.[55, 74, 75]
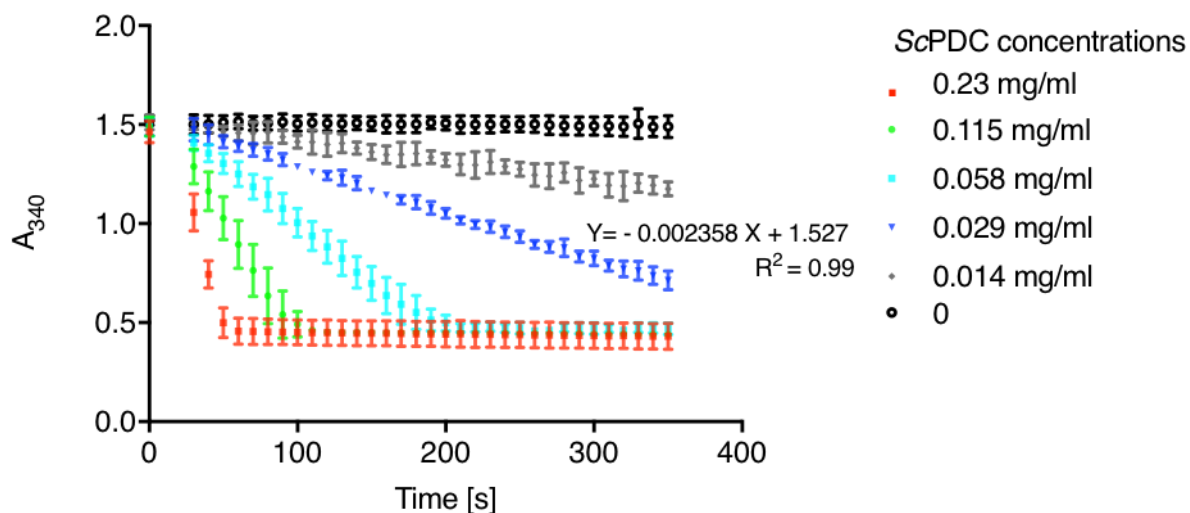
Figure 3-14. *Sc*PDC decarboxylase activity test by coupled assay at varying concentrations of enzyme

3.2.5. Effect of pH on the decarboxylase activity of *Sc*PDC

The decarboxylase activity *Sc*PDC is known to be optimum in the pH range of 6.0- 6.8, in both citrate and MES buffers.[73,76] The pH studies on the decarboxylation reaction of *Sc*PDC is limited to above buffers. Therefore, exploring the pH dependency of *Sc*PDC catalysed reaction is of our interest in order to understand the effect of pH on the decarboxylation reaction of the enzyme. This is important to see whether decarboxylation and carboxylation reactions which will be examined later in this thesis can occur at the same pH. For this purpose, instead of citrate, three other buffers were employed. These included 200 mM sodium phosphate (NaH_2PO_4-Na_2HPO_4, pH 6.0-8.0), 200 mM HEPES (pH 7.5-8.5) and 200 mM sodium carbonate ($NaHCO_3$-Na_2CO_3, pH 8.0-9.0).

Generally, *Sc*PDC exhibited higher activity in HEPES buffer compared to phosphate buffer at the same pH values (pHs 7.5 and 8.0). This can be explained by the inhibitory role of phosphate, as it mimics the structure of the pyruvate.[77] As illustrated in Figure 3-15, no decarboxylase activity was observed at higher pH values. In fact, the enzyme lost its decarboxylase activity at pH of 8.5 in all three buffers. This observation was significant as the reverse reaction of *Sc*PDC, carboxylation, was previously reported in the sodium carbonate buffer (100 mM) at pH 8.5 and above (pH 8.5-11.5).[45]
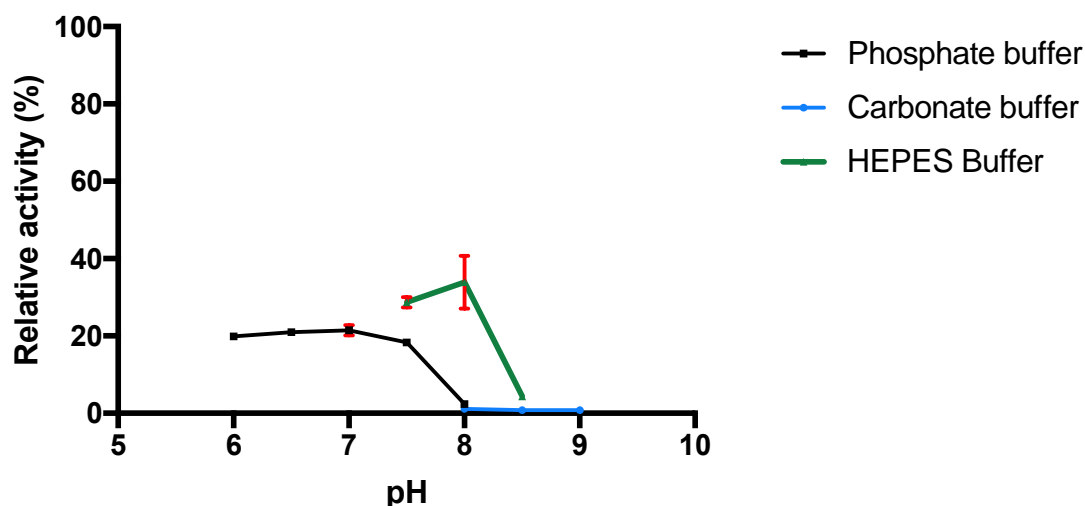
Figure 3-15. Relative decarboxylase activity (%) of *Sc*PDC in various buffers in comparison with citrate. The measurements were performed by the coupled assay at 30 °C. Error bars in red indicate standard deviations (SD) for three repeat experiments.

In order to establish that the loss in the decarboxylase activity of *Sc*PDC in the carbonate buffer was independent from the inactivation of the coupling enzyme, the activity of *Sc*ADH was measured independently in two different buffers. These included 200 mM sodium carbonate (NaHCO$_3$-Na$_2$CO$_3$, pH 8.0-9.5) and 200 mM sodium phosphate (NaH$_2$PO$_4$-Na$_2$HPO$_4$, pH 6.0-7.5).
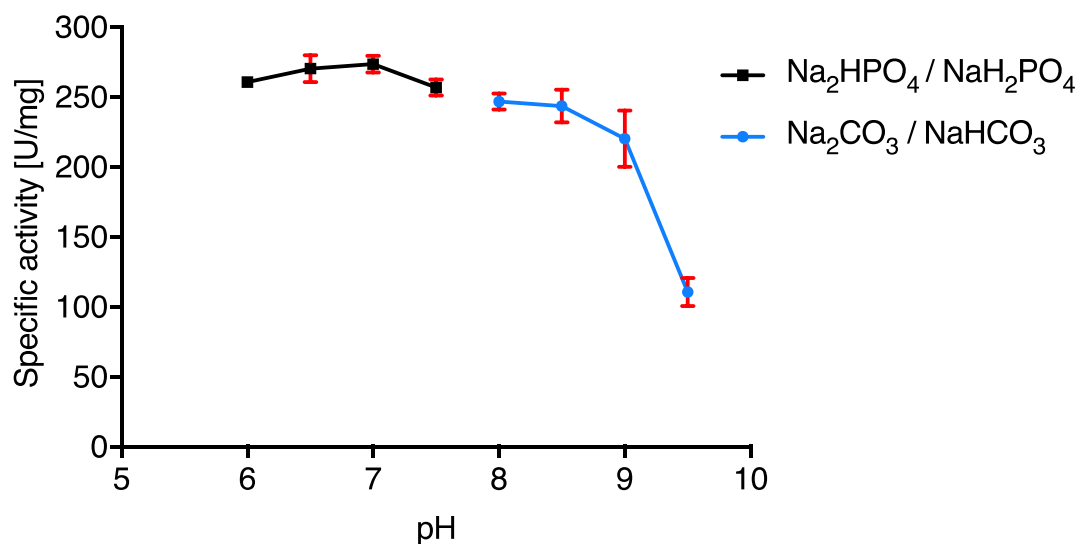
Figure 3-16. Determination of the activity of *Sc*ADH in sodium carbonate and sodium phosphate buffers. Error bars in red indicate standard deviations (SD) for three repeat experiments.

As illustrated in Figure 3-16, no loss in the oxidative activity of ScADH was observed at pH 8.5 of the carbonate buffer. In fact, the enzyme activity was almost remained constant until the pH increased from 9.0 to pH 9.5. If ScADH shows similar results in the reduction direction, then it could be concluded that the loss in the decarboxylase activity of the ScPDC in the carbonate buffer could be due to the inactivation of the enzyme itself. However, this is not the case as the reductive activity of ScADH has been reported to be preferred at lower pHs. This is due to the protonation of the aldehyde substrate at lower pHs which facilitates the reduction.[78]

3.2.6. Effect of pH on the stability of ScPDC

The carboxylation reaction of ScPDC is relatively slow, being the reverse of the reaction, the enzyme has evolved to achieve. According to the available thermodynamic data,[79] the reverse reaction of ScPDC, formation of pyruvate from acetaldehyde and CO_2, has been reported to be extremely unfavourable.[80] After overcoming the unfavourable equilibrium by different methods, like coupling to a favourable reaction, the yield may be improved by increasing the reaction time. However, increasing the incubation period would affect the stability of the enzyme.

Therefore, it was decided to investigate the stability of ScPDC over a specific period in different buffers with various pHs. The pH stability of the ScPDC was monitored in three different buffers having various pH values at three intervals (0, 24 and 48 hr). The enzyme was incubated in sodium phosphate (200 mM, pH 6.0-7.5), HEPES (200 mM, pH 7.5-8.5) and sodium carbonate buffer (200 mM, pH 8-9.5) at 25 ℃. After the incubation period, the decarboxylase activity of the enzyme was measured by the coupled assay. The activity of the ScPDC in both phosphate and HEPES buffer slightly decreased over time. On the other hand, the enzyme was unstable in the carbonate buffer and lost its decarboxylase activity rapidly with no detectable activity after 48 hours (Figure 3-17).
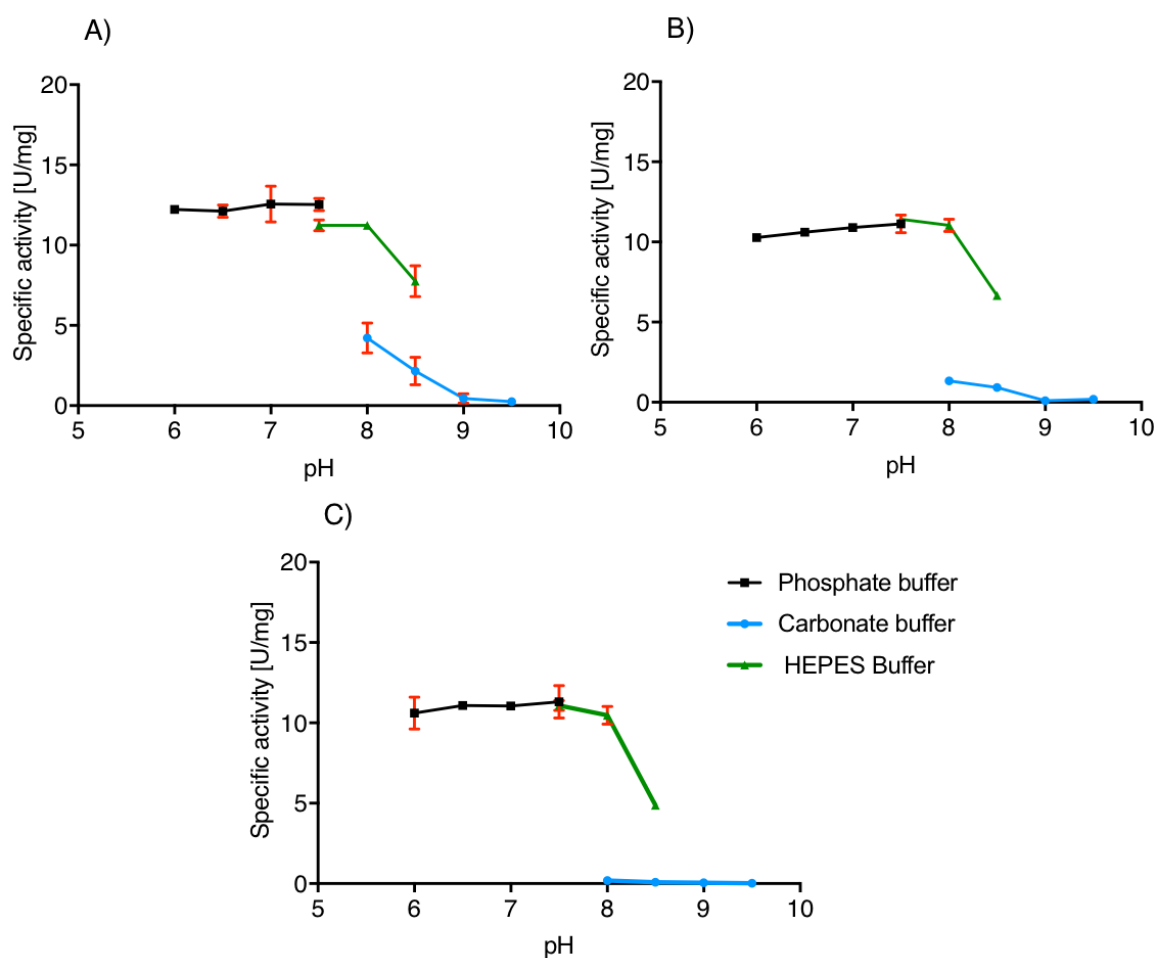
Figure 3-17. Effect of pH on the stability of *Sc*PDC at different intervals. A) time zero B) 24 hours and C) 48 hours. Measurements were made by the coupled assay. Error bars in red indicate standard deviations (SD) for three repeat experiments.

3.2.7. Carboxylation of different aldehydes using the reverse reaction of *Sc*PDC under atmospheric pressure

Next, the reverse reaction of the enzyme was investigated in order to explore the CO_2 fixation property of this enzyme. For this purpose, the carboxylase activity of the enzyme was examined with two different aliphatic aldehydes: acetaldehyde and propanal. These aldehydes were selected as their corresponding keto acids are pyruvate and 2-ketobutyric acid. Pyruvate is the natural substrate of *Sc*PDC and 2-ketobutyric acid is the second closest structure to pyruvate. Two different experiments were set up separately. Since the carboxylase activity of the *Sc*PDC was reported in the sodium carbonate buffer (pH≥8.5) for the first time, here the same buffer was selected.[45] The reaction was performed using carbonate buffer (100mM, pH 8.5) as the source of CO_2 under atmospheric pressure (78%

nitrogen, 21% oxygen, 1% others). Acetaldehyde or propanal were added to the reaction at final concentration of 0.1 mM (same concentration as used by Miyazaki *et al.*).[45] After 1 hour the reaction was analysed by HPLC for the detection of α-ketoacids. Neither of the products, namely pyruvate nor 2-ketobutyric acid were detected. The concentration of starting aldehyde substrates was increased from 0.1 mM to 10 mM. Again, the reactions failed to produce any α-ketoacids. Next, the reaction time was increased from 1 hour to 48 hours to see whether increasing the time of the reaction makes any changes. No product formation was observed. The results obtained here were in contrast with the published data in which the carboxylation reaction of acetaldehyde to pyruvate catalysed by *Sc*PCD was reported in carbonate buffer.[45] These results demonstrate that the carbonate buffer alone is not a sufficient source of CO_2 to achieve the reverse the reaction of *Sc*PDC. In fact, saturated concentration of CO_2 is required in order to shift the position of the equilibrium of the *Sc*PDC catalysed reaction towards unfavoured carboxylation, as determined by Le Chatelier's principle. This can be achieved by performing the reaction under a pressurised CO_2 environment, a topic which will be explored later in this thesis.

3.1. Conclusion:

In this chapter, *Sc*PDC gene was successfully amplified and cloned into pJexpress414 vector. The enzyme was expressed having a N-terminal His-tag which facilitates the purification step. Both activity and stability of the purified enzyme were studied in buffers at different pHs. This was done by a coupled assay using ADH as the coupling enzyme. The activity of ADH was also established in those buffers in order to make sure that the changes in the results is not influenced by the changes in activity of ADH. Here, the decarboxylase activity of *Sc*PDC was completely abolished in carbonate buffer. This is because *Sc*PDC is less stable at pH 8.0 or above.

On the other hand, the carboxylation reaction of *Sc*PDC was reported in carbonate buffer, with improvement in the yield of reaction by increasing the pH of the carbonate buffer (from pH 8.5 to 11.5).[45] This is due to the reaction between the product of the carboxylation reaction, 2-keto acid, with the base resulting in the formation of water and the salt of 2-keto acid. Although no decarboxylase activity was observed in carbonate buffer, the reverse reaction of the enzyme was examined in this buffer as tested and reported by Miyazaki *et*

al.[45] Our results showed that *Sc*PDC is unable to catalyse the carboxylation reaction using carbonate buffer as the source of CO_2. In the next chapter a second thiamine diphosphate decarboxylase enzyme, a branched chain decarboxylase will be investigated for its potential application for the CO_2 fixation reaction using carbonate buffer as a source of CO_2.

Chapter 4. Preparation of a thiamine dependent decarboxylase enzyme branched chain decarboxylase from *Lactococcus lactis*

This chapter gives a summary highlighting branched chain decarboxylase (KdcA) and its structure. A synthetic gene of *kdcA* from *Lactococcus lactis* (*Ll*KdcA) in a pET32a vector has been expressed and purified. The activity of purified *Ll*KdcA has been tested in the decarboxylation direction with different 2-keto acids. Finally, the reverse reaction (carboxylation) of this enzyme has been examined using carbonate buffer as the only source of CO_2.

4.1. Background

α-Ketoacid decarboxylases are involved in the degradation of 2-ketoacids. They occur more frequently in plants, fungi and yeasts compared to bacteria.[81] In 2004, a novel branched chain 2-keto acid decarboxylase (KdcA) from *Lactococcus lactis* (*Ll*KdcA) was identified as an α-ketoacid decarboxylase enzyme.[81] The action of the enzyme was described to be important in flavour formation during cheese ripening process. Later, sequence comparison of *Ll*KdcA with three other ThDP-dependent decarboxylases (*Sc*PDC, pyruvate decarboxylase from *Zymomonas mobilis: Zm*PDC and indolepyruvate decarboxylase from *Enterobacter cloacae*: *Ec*IPDC) suggested that the enzyme is a new member of the ThDP-dependent decarboxylase family.[82] In comparison to other decarboxylases, *Ll*KdcA was shown to have a broad substrate spectrum with the optimum activity against 3-methyl-2-oxobutanoic acid (α-ketoisovaleric acid).[82] For this reason the enzyme is also called α-ketoisovalarate decarboxylase (Kivd). *Ll*KdcA catalyses the decarboxylation of 3-methyl-2oxobutanoic acid to isobutyraldehyde and CO_2 (Scheme 4-1).
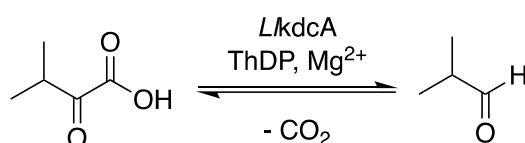
Scheme 4-1. Decarboxylation of 3-methyl-2oxobutanoic acid to isobutyraldehyde by *Ll*KdcA

The mechanism of decarboxylation reaction catalysed by *Ll*KdcA follows the same reaction steps as described in section 3.1.2 for *Sc*PDC and involves the formation of thiazolium ylide as

the crucial reaction intermediate. In 2006, the structure of *Ll*KdcA was modelled based on the crystal structures of three other decarboxylase enzymes: *Zm*PDC, *Sc*PDC and *Ec*IPDC.[83] From 32 structures generated for *Ll*KdcA, the final model represented the enzyme to be a dimer.[83] This was in contrast with *Sc*PDC with a tetrameric structure. One year later, the crystal structure of recombinant *Ll*KdcA was determined at 1.6 and 1.8 Å resolution, confirming the homo-dimeric structure of *Ll*KdcA (Figure 4-1).[84] Each subunit consists of 547 amino acids with the α/β topology observed in all ThDP-dependent decarboxylase enzyme. There are three domains within each monomer: α-domain, β-domain and γ-domain. (Figure 4-1).
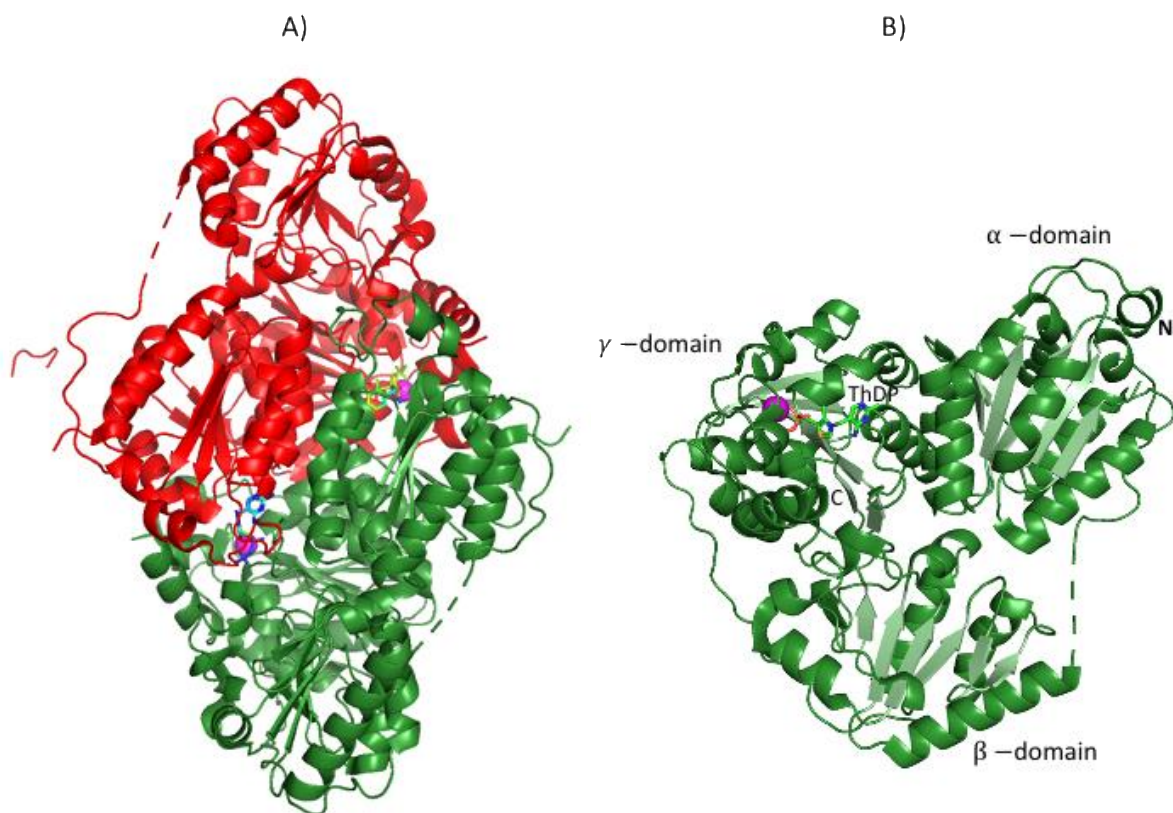
Figure 4-1. Crystal structure of *Ll*KdcA (PDBID:2VBF): (A) homodimer (B) monomer consisting of three domains within a subunit: α, β and γ domain. PYR, PP and R domain. The ThDP is shown in stick and magnesium ion in sphere. The figure was created by PyMOL.

The active site is located at the interface PYR and PP domains of two subunits. Hence two monomers are required to form catalytically active enzyme.[84] The typical V conformation of the ThDP is also found in *Ll*KdcA which is stabilized by the a large hydrophobic residue (Ile404).[85]

4.2. Results and discussion

4.2.1. Expression and purification of *Ll*KdcA

The *kdcA* gene from *L. lactis* (Uniprot ID: Q6QBS4) was codon-optimized for expression in *E. coli* and synthesised by DC Bioscience Ltd, UK. The *kdcA* gene was subsequently cloned into the pET32a expression vector (DC Bioscience Ltd, UK). DNA sequencing confirmed the correct identity of the *kdcA* sequence as it was ordered. The recombinant *Ll*KdcA was designed with a hexa histidine amino acid tag (His-tag) followed by a TEV cleavage site at the N-terminal. The incorporation of the His-tag in the sequence was to facilitate the protein purification via the IMAC in a single step. Here, each subunit of *Ll*KdcA was extended from 547 amino acids to 568 because of the 21-amino acid extension at the N-terminal (Figure 4-2).

```
         10         20         30         40         50         60
MGSSHHHHHH SSGENLYFQG SMYTVGDYLL DRLHELGIEE IFGVPGDYNL QFLDQIISRE

         70         80         90        100        110        120
DMKWIGNANE LNASYMADGY ARTKKAAAFL TTFGVGELSA INGLAGSYAE NLPVVEIVGS

        130        140        150        160        170        180
PTSKVQNDGK FVHHTLADGD FKHFMKMHEP VTAARTLLTA ENATYEIDRV LSQLLKERKP

        190        200        210        220        230        240
VYINLPVDVA AAKAEKPALS LEKESSTTNT TEQVILSKIE ESLKNAQKPV VIAGHEVISF

        250        260        270        280        290        300
GLEKTVTQFV SETKLPITTL NFGKSAVDES LPSFLGIYNG KLSEISLKNF VESADFILML

        310        320        330        340        350        360
GVKLTDSSTG AFTHHLDENK MISLNIDEGI IFNKVVEDFD FRAVVSSLSE LKGIEYEGQY

        370        380        390        400        410        420
IDKQYEEFIP SSAPLSQDRL WQAVESLTQS NETIVAEQGT SFFGASTIFL KSNSRFIGQP

        430        440        450        460        470        480
LWGSIGYTFP AALGSQIADK ESRHLLFIGD GSLQLTVQEL GLSIREKLNP ICFIINNDGY

        490        500        510        520        530        540
TVEREIHGPT QSYNDIPMWN YSKLPETFGA TEDRVVSKIV RTENEFVSVM KEAQADVNRM

        550        560
YWIELVLEKE DAPKLLKKMG KLFAEQNK
```

Figure 4-2. Gene design and primary sequence of recombinant *Ll*KdcA. His-Tag in red, TEV site in green and *Ll*KdcA in blue

General properties of *Ll*KdcA were estimated based on its amino acid sequence using ProtParam software as summarized in (Table 4-1).

Table 4-1.General properties of a single subunit of *Ll*KdcA.Data generated using ProtParam software (ExPASY Protemics tools)

	Length (aa)	Molecular weight (kDa)	Theoretical pI	Extinction coefficient at 280 nm ($M^{-1}cm^{-1}$)
*Ll*KdcA	568	63.24	5.12	54320

4.2.2. Recombinant over-expression of *Ll*KdcA in *E. coli*

Sequenced-confirmed *Ll*KdcA-pET32a expression vector was used to transform the BL21 *E. coli* expression host. Initial expression tests were performed on a small scale (20 mL cultures) in order to evaluate the extent of induction and solubility of *Ll*KdcA when expressed as a recombinant protein in *E. coli.* Recombinant *Ll*KdcA was expressed as a His-tag protein following the addition of 1mM IPTG,[83] and the growth of cells at 15 °C or 20 °C for 18 hours. Theoretically, the growth of cells at lower temperature following the induction would increase the expression of soluble protein. After overnight growth, bacterial cells were harvested and lysed chemically. Soluble and insoluble fractions were separated. All fractions including non-induced, induced soluble and induced insoluble were analysed by 12% (w/v) SDS-PAGE gel electrophoresis (Figure 4-3).

Protein bands were detected by Coomassie blue staining. Bands of expected size (63.24 kDa) corresponding to His-tagged version of the protein were present in both soluble and insoluble fraction when IPTG was added. However, there was slightly less insoluble protein when the cells were grown at 15 °C following the induction.
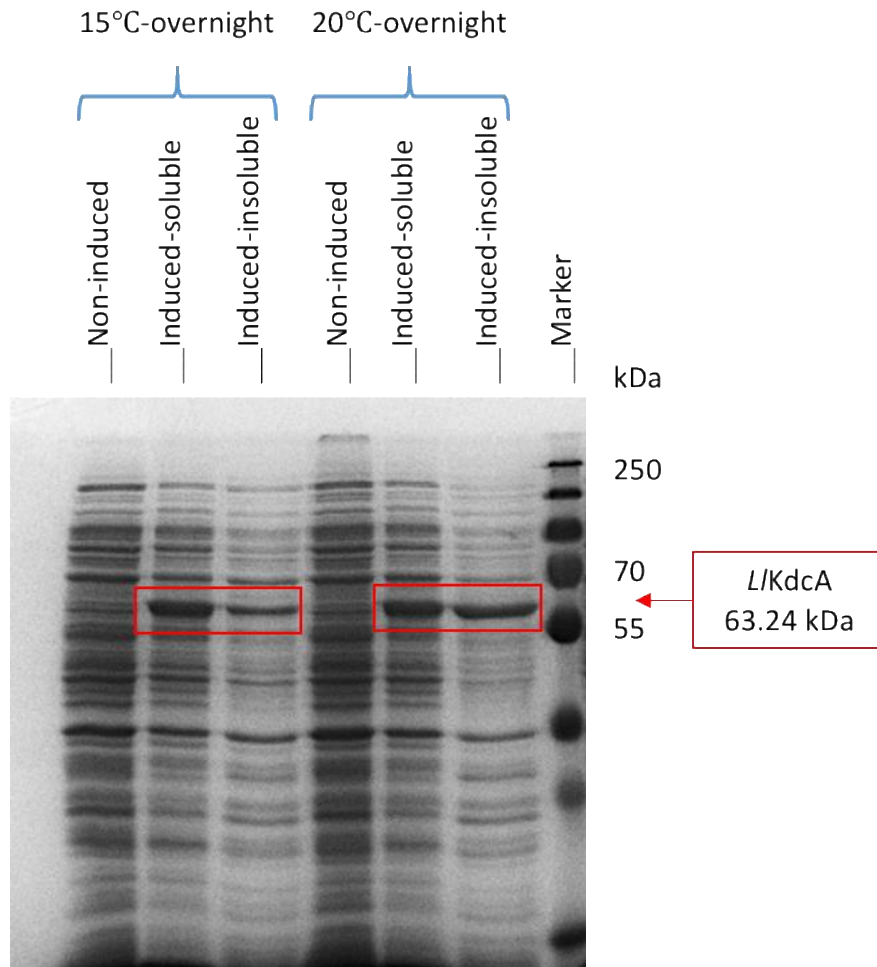
Figure 4-3. SDS-PAGE (12%) analysis of *Ll*KdcA expression at different temperatures. Marker is PageRuler™ Plus Prestained Protein Ladder from ThermoFisher.

4.2.3. Purification of recombinant *Ll*KdcA

The recombinant (His)$_6$-*Ll*KdcA protein was expressed in 1 litre LB culture. Once the OD$_{595}$ reached 0.6, the cells were induced with IPTG (1mM) and grown further at 15 °C for 18 hours. Bacterial cell pellets containing *Ll*KdcA were harvested and stored at – 80 °C until needed. The composition of the buffers used to purify *Ll*KdcA was similar to those used for *Sc*PDC. However, sodium phosphate (50 mM, pH 6.5) was used instead of citrate. After sonication, the cell free extract was applied to a Ni-NTA column and the column was washed successively with 5% and 10% buffer B. (His)$_6$-*Ll*KdcA bound to the column was eluted with 30% buffer B. All fractions were analysed by 12% (w/v) SDS-PAGE (Figure 4-4).
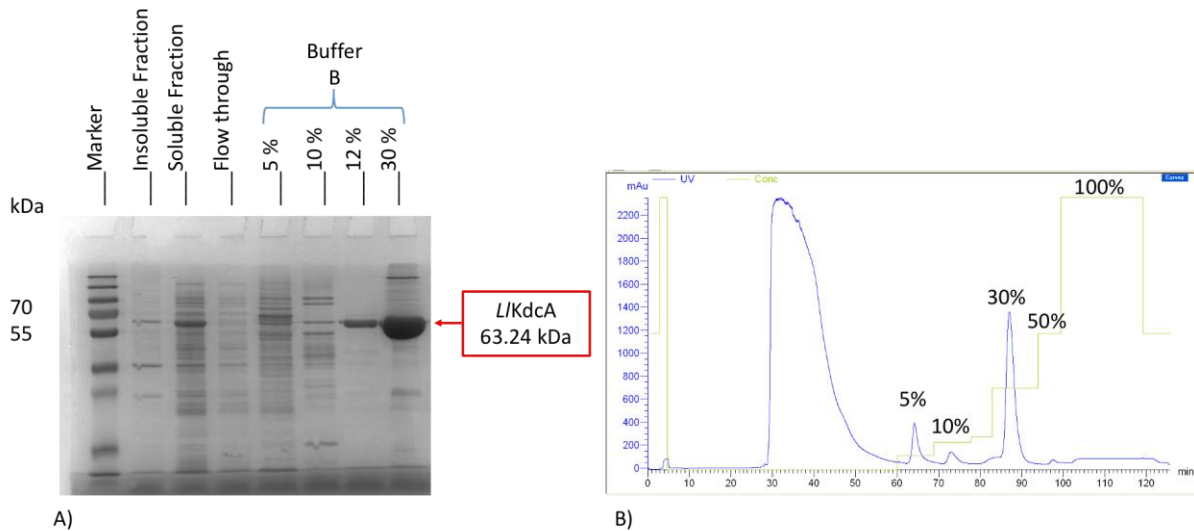
Figure 4-4. Expression and purification of (His)$_6$-*Ll*KdcA protein using Ni^{2+} affinity chromatography and AKTA technology. (A) Comassie stained SDS-PAGE (12%) of the proteins present in different fractions including insoluble, soluble, flow through, wash steps and finally elution of *Ll*KdcA with different concentrations of imidazole and Marker (PageRuler™ Prestained Protein Ladder) (B) Representative chromatogram for purification of (His)$_6$-*Ll*KdcA recombinant protein using Ni^{2+} affinity chromatography and AKTA technology. Absorbance at 280 nm is shown in blue and the concentration of the buffer B used in green.

Fractions containing (His)$_6$-*Ll*KdcA showed a band between 70 kDa and 55 kDa. These fractions (12% and 30% buffer B) were pooled and subsequently dialysed (MWCO 3500 Da) overnight at 4 °C against buffer C. After dialysis the sample was concentrated to 11 mg/ml using a vivaspin column (100 kDa MWCO).

4.2.4. *Ll*KdcA decarboxylase activity tested by coupled assay

The substrate range of *Ll*KdcA is of interest in order to check the activity of the purified enzyme in the decarboxylation direction. This would also provide an insight into the possible aldehydes that could be the substrate of the enzyme in the CO$_2$-fixation direction. Here, the decarboxylase activity of the purified *Ll*KdcA was measured by the standard coupled assay (Section 2.7) in the sodium citrate buffer (200 mM, pH 6.0) at 30 °C with the following substrates: pyruvate (**1**), 2-ketobutanoic acid (**2**), 2-ketopentanoic acid (**3**), 4-methyl-2-oxopentanoic acid (**4**) and 3-methyl-2-oxobutanoic acid (**5**) at a final concentration of 30 mM.
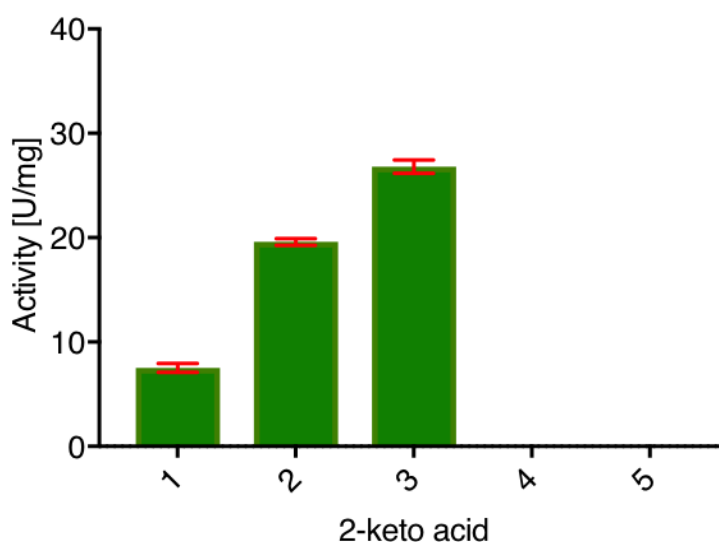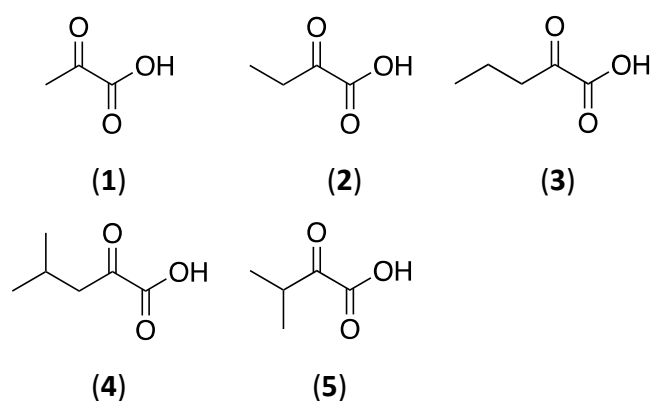
Figure 4-5. Substrate range of *Ll*KdcA with various 2-keto acids in the decarboxylation direction. Coupled assay was used to measure the specific activity. Assay condition: Sodium citrate buffer (200 mM, pH 6.0) at 30 °C with each substrate at the final concentration of 31.25 mM. Error bars in red indicate standard deviations (SD) for three repeat experiments.

In contrast to *Sc*PDC, the decarboxylase activity of the *Ll*KdcA gradually increased when the length of the side chain of the 2-keto acid substrate increased from one to three carbons (Figure 4-5). This observation was comparable with the reported literature values for the decarboxylase activity of a purified *Ll*KdcA containing an N-terminal (His)$_6$ tag.[86] However, no activity was recorded for the branched chain substrates (**4**) and (**5**) using the standard coupled assay. This is because the coupling enzyme *Sc*ADH has no activity with the generated aldehydes from both (**4**) and (**5**).[78] Also, it was reported that the aldehydes generated from decarboxylation of (**4**) and (**5**) interfere with the coupled assay due to their absorption at the

wavelength of 340nm.[86] Therefore, a direct decarboxylase assay as developed by Gocke *et al.* was followed to measure the activity of *Ll*KdcA.[86] In this assay the decomposition of substrate (**4**) and (**5**) were followed directly at 340 nm by adding the enzyme to each substrate separately.[86] Prior to the activity measurements the extinction coefficient of both (**4**) and (**5**) were obtained in sodium citrate buffer (200 mM, pH 6.0) at 30 ℃ (Figure 4-6).
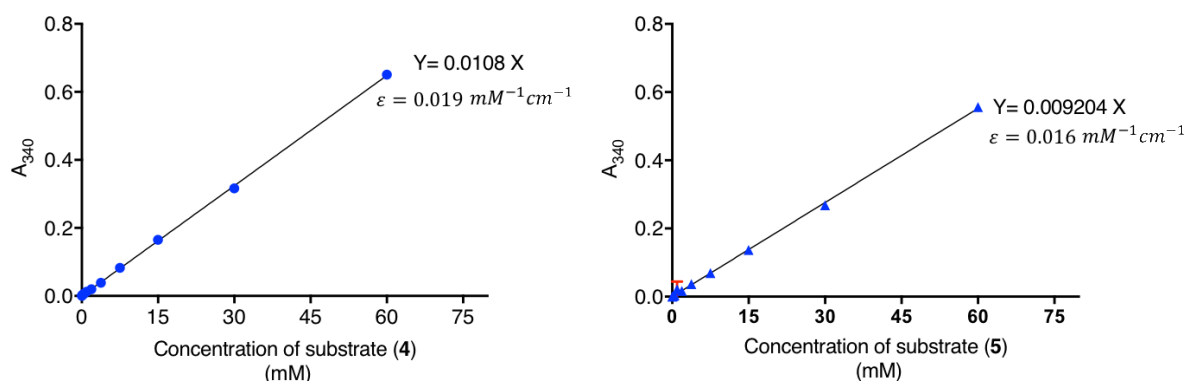
Figure 4-6. Determination of extinction coefficient of substrates (**4**) and (**5**) in sodium citrate buffer (200 mM, pH 6.0) at 30 ℃. Error bars in red indicate standard deviations (SD) for three repeat experiments.

The activity measurements of *Ll*KdcA using the direct assay showed the enzyme to have higher decarboxylase activity with branched chain substrates including (**4**) and (**5**). The maximum activity was observed with substrate (**5**) being the natural substrate of the enzyme (Figure 4-7). Generally, a broader substrate range was observed for *Ll*KdcA compared to *Sc*PDC as reported elsewhere.[82]
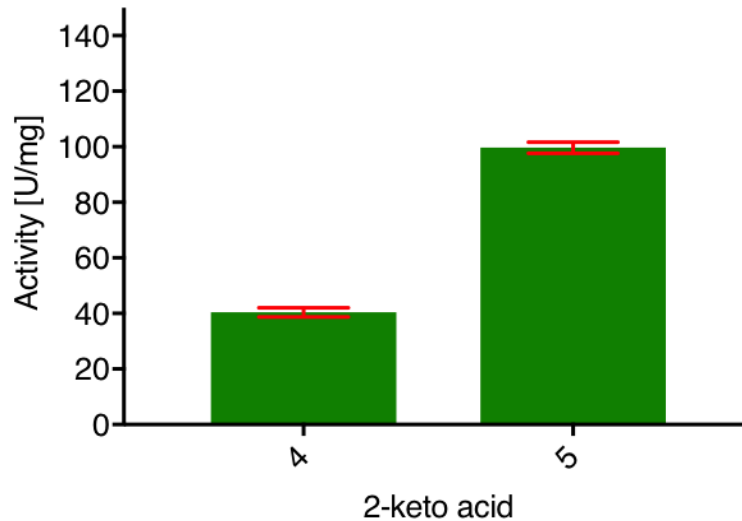
Figure 4-7. Decarboxylase activity of *Ll*KdcA with substrates (**4**) and (**5**) measured by the direct assay. Assay condition: Sodium citrate buffer (2004.2.4 mM, pH 6.0) at 30 ℃. Error bars in red indicate standard deviations (SD) for three repeat experiments.

4.2.5. Effect of pH on the decarboxylase activity of *Ll*KdcA

In order to understand the effect of pH on the activity of *Ll*KdcA and compare it with the results obtained for *Sc*PDC in section 3.2.5, the decarboxylase activity of this enzyme was measured at three buffers having different pH values. These included 200 mM sodium phosphate (NaH_2PO_4-Na_2HPO_4, pH 6.0-8.5), 200 mM HEPES (pH 7.5-8.5) and 200 mM sodium carbonate ($NaHCO_3$-Na_2CO_3, pH 8.0-9.5).
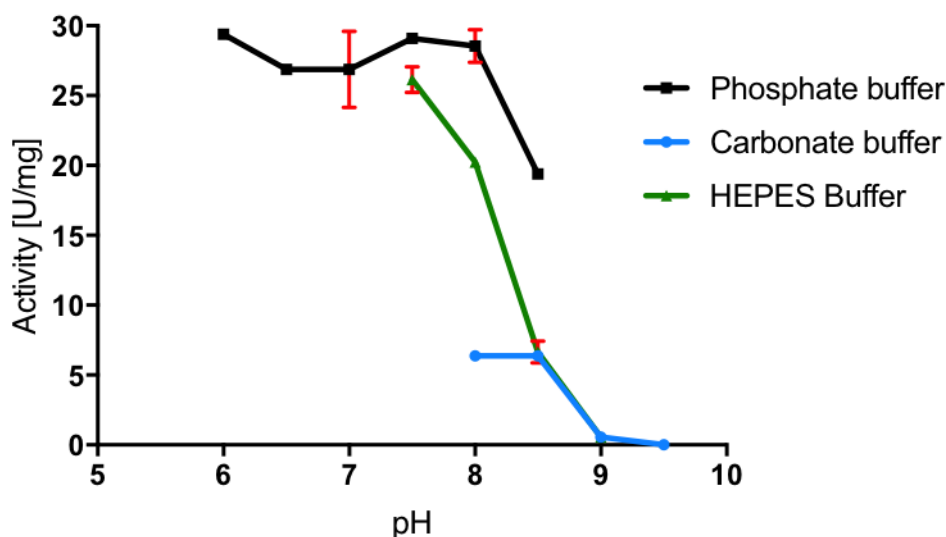
Figure 4-8. Effect of pH and buffer on the *Ll*KdcA decarboxylase activity. The measurements were determined by the coupled assay at 30 ℃ in 200 mM phosphate (square), HEPES (triangle) and carbonate buffer (circle) with 2-oxovaleric acid (**3**) as the substrate. Error bars in red indicate standard deviations (SD) for three repeat experiments.

As shown in Figure 4-8, the decarboxylase activity of the enzyme was optimum at pH 6.0 of the phosphate buffer (200 mM). This activity was similar to that observed in the standard buffer, citrate (200 mM) at pH 6.0 in section 4.2.4 , showing that the activity of the enzyme is not inhibited by the phosphate. This is in contrast to *Sc*PDC where its decarboxylase activity is less (23%) in the phosphate buffer compared to the citrate buffer at the same pH value (pH 6.0). Increasing the pH of phosphate buffer from 6.0 to 7.5 did not have a significant effect on the activity of the enzyme. At pH 8.0 the enzyme was still active in all three buffers with the highest activity in the phosphate followed by HEPES and then carbonate. This is in contrast to *Sc*PDC where no decarboxylase activity was detected at pH 8.0 in either phosphate or carbonate buffers (section 3.2.5). In 2018, the carboxylation reaction of *Ll*KdcA in combination with an aminotransferase enzyme was exploited for the synthesis of L-methionine from methional and CO_2.[29] In this example, the experiments were conducted in carbonate buffer (200 mM, pH 8.0). Our results showed that the decarboxylase activity of the enzyme was significantly lower in this buffer. A similar observation was made for *Sc*PDC as reported in the previous chapter.

4.2.1. Carboxylation reaction of *Ll*KdcA

Here, the investigation of the carboxylase activity of *Ll*KdcA is of interest in order to understand whether the ThDP-dependent decarboxylase enzymes such as *Sc*PDC and *Ll*KdcA are able to catalyse the carboxylation reaction on their own in the presence of carbonate buffer. This is the first time where the carboxylase activity of *Ll*KdcA is tested in the carbonate buffer as the only source of CO_2 in a single reaction (not coupled in a cascade). To test the CO_2 fixation activity of *Ll*KdcA, the carboxylase activity of the enzyme was examined with propanal as the substrate. Propanal was selected as it is the product of the decarboxylation of 2-ketobutanoic acid (**2**). This substrate is accepted by the enzyme in the decarboxylation direction. Also, its higher boiling point compared to acetaldehyde makes it less volatile and easier to quantify at room temperature. The reaction was setup as explained for *Sc*PDC using sodium carbonate buffer (100mM, pH 8.5) as the source of CO_2 under atmospheric pressure. Propanal were added to the reaction at final concentration of 25 mM. After 1 hour, the reaction was analysed by HPLC for the detection of α-ketoacids. No 2-ketobutyric acid was detected. The reaction incubation time was increased from 1 hour to 48 hours to see whether that improves the reaction. Again, no product formation was observed. These results are similar to the one that was obtained with *Sc*PDC, showing no carboxylase activity in the carbonate buffer. Evidently, carbonate buffer (100 mM) is not a sufficient source of CO_2 in order to achieve the carboxylation of aliphatic aldehydes using the reverse reaction of *Sc*PDC and *Ll*KdcA.

4.3. Conclusions

In this chapter, *Ll*KdcA was expressed and purified as a His-tag protein. The decarboxylase activity of the enzyme was tested in various buffers. The enzyme was shown to have a broader pH profile compared to *Sc*PDC. Unlike *Sc*PDC where its decarboxylase activity was lost in the carbonate buffer, *Ll*KdcA showed to still have decarboxylase activity in this buffer (until pH 8.5). However, the carboxylation reaction of *Ll*KdcA in the carbonate buffer failed to generate any α-keto acid. In summary, the conversion of CO_2 to useful chemicals using the reverse reaction of decarboxylase enzymes, including *Sc*PDC and *Ll*KdcA, is a challenging process. It requires the presence of an adequate concentration of CO_2 in order to shift the position of the equilibrium towards the unfavoured carboxylation. Also, coupling the carboxylation

reaction to other enzymatic reaction(s) can be used as a strategy to favour the carboxylation reaction. In the next chapter, aminotransferase enzymes will be investigated as potential enzymes to be coupled with the carboxylation reaction of both *Sc*PDC and *Ll*KdcA.

Chapter 5. Transaminases as auxiliary enzymes to drive carboxylation

5.1. Background

5.1.1. Transaminases

In nature, transaminases (TAs) or aminotransferases (ATs) (EC 2.6.1.X), are the most important enzymes for the synthesis and catabolism of α-chiral amino acids and amines. In the overall reaction, transaminases catalyse the reversible transfer of an amino group from an amine donor to a carbonyl carbon atom acceptor molecule (α-keto acid, aldehyde or ketone)[87] (Scheme 5-1).
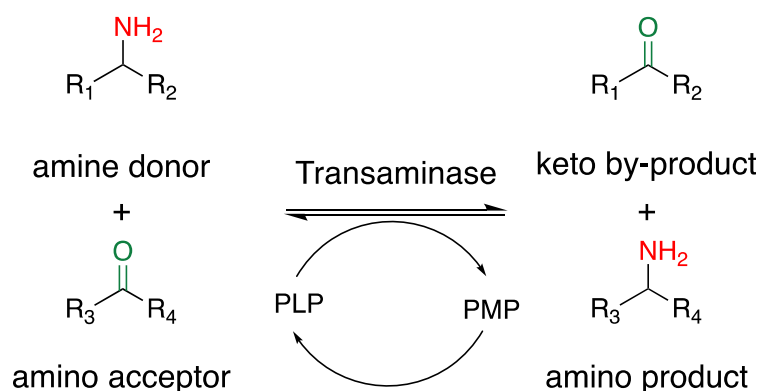
Scheme 5-1. Overall transaminase catalysed reaction. Pyridoxal-5'-phosphate (PLP), pyridoxamine-5'-phosphate (PMP).

The transaminase apoenzyme requires pyridoxal-5'-phoshpate (PLP), a derivative of vitamin B_6, as the coenzyme to generate its holoenzyme (Scheme 5-2).[88]
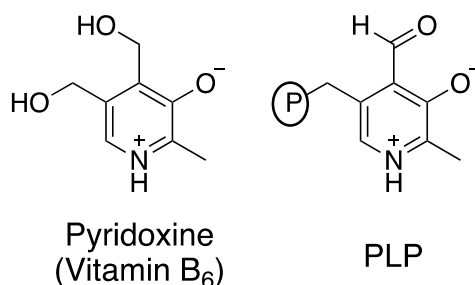
Scheme 5-2. Chemical structure of vitamin B_6 and PLP. P represents PO_4^{2-}.

The coenzyme is bound reversibly to the enzyme through a Schiff-base linkage to the ε-amino group of an active site lysine residue.

5.1.2. Mechanism of transamination reaction

The mechanism of a transaminase catalysed reaction follows a *ping-pong bi-bi* mechanism. The main characteristic of the ping-pong mechanism is that the enzyme reacts with the first substrate to generate a product which must dissociate before the second substrate can enter the active site of the enzyme to bind to a modified enzyme or cofactor.[89] There are two half-reactions involved in the overall mechanism of transamination (Scheme 5-3).
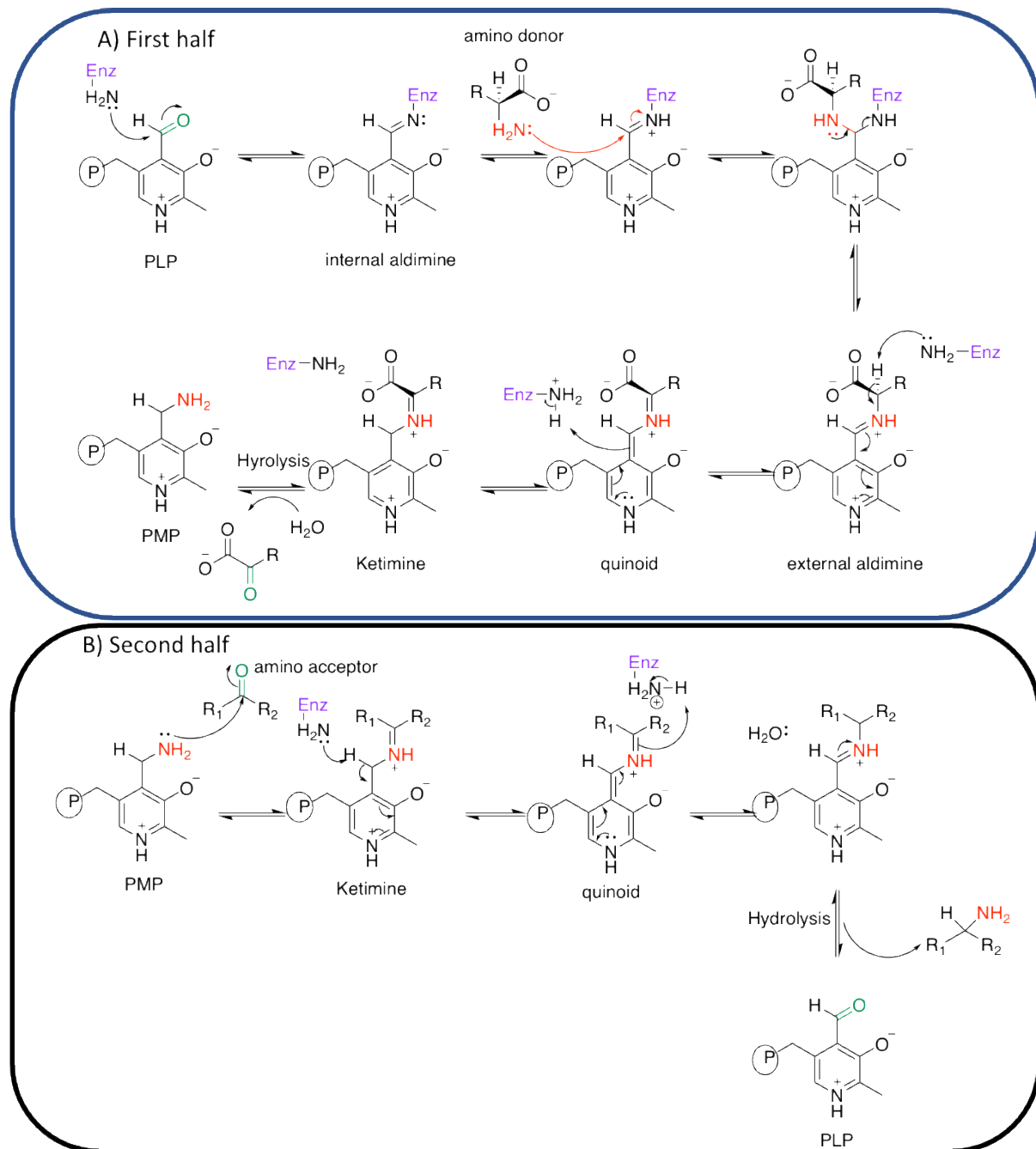
Scheme 5-3. Mechanism of transamination reaction which involves two half-reactions.[90] Pyridoxal-5'-phosphate (PLP), pyridoxamine-5'-phosphate (PMP) and PO_4^{2-} (P).

First half: In the active site of the enzyme, PLP is bound to the enzyme via an imine linkage between the aldehyde group of the cofactor PLP and the ε-amino group of a conserved lysine residue forming the internal aldimine intermediate. This intermediate reacts with an amine donor substrate and forms an external aldimine. The pyridine ring of the external aldimine acts as an electron sink which results in the formation of a ketamine through a quinoid intermediate. The ketamine intermediate undergoes a hydrolysis reaction which subsequently generates pyridoxamine-5'-phosphate (PMP) which is retained at the active site and the carbonyl containing by-product is released (Scheme 5-3, A). Once PMP is generated from the first half, the second half of the reaction will proceed.

Second half: The PMP generated from the first half of the reaction acts as an amine donor and reacts with the second substrate (amine acceptor) of the reaction. The second half of the reaction follows the exact sequence of reaction events of the first half in the reverse direction which ultimately regenerates PLP (Scheme 5-3, B). In contrast to PLP which covalently binds to the enzyme, PMP stays in the active site by non-covalent interactions.[91]

5.1.3. Classification of TAs according to their amino donor

Transaminases (TAs) are very diverse enzymes in terms of their substrate specificity. According to the BRENDA-database there are about 81 distinct enzymes from this family. TAs can be classified based on their amino donor, particularly in terms of the presence and the position of the amino group being transferred, relative to a negatively charged group (carboxyl group). In principle, there are three different groups: alpha transaminase (α-TAs), omega transaminases (ω-TAs) and amine transaminases (Scheme 5-4).[92]
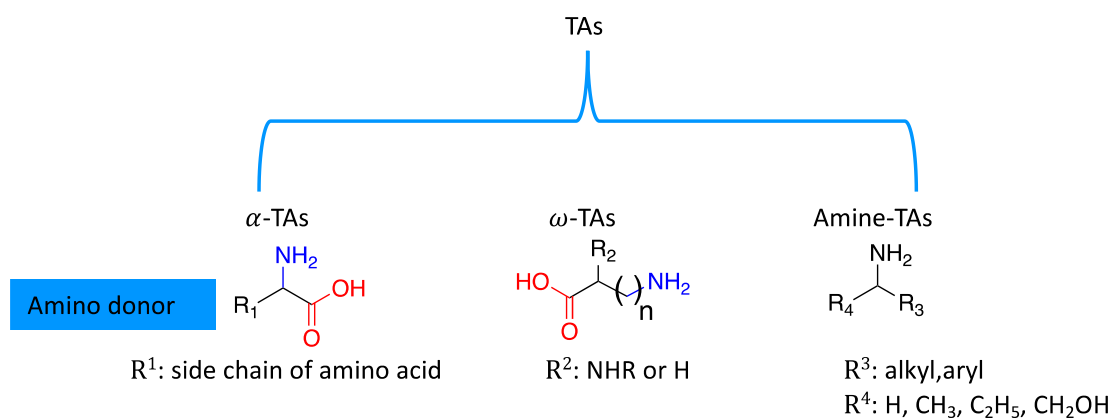
Scheme 5-4. Classification of TAs based on the amino donor.

The α-TAs catalyse the transfer of the amino group located at the α carbon relative to the carboxyl group. These enzymes usually catalyse the transamination of natural α-amino acids in metabolism. The ω-TAs belong to a smaller group of the TA family and catalyse the transfer of a terminal amino group located further away from the carboxyl group (at least two carbons away from a carboxyl group). Finally, amine transaminases have no carboxyl group in the amino donor.[87]

TAs have also been divided into six subgroups according to their sequence similarities and secondary structure predictions (class I, class II, class III, class IV, class, V and class VI).[93, 94] Class I and II includes L-aspartate transaminase and L-alanine transaminase, class III ω-transaminase, class VI D-amino acid transaminase and branched chain transaminase, class V L-serine transaminase and class VI sugar transaminase.

5.2. Application of TAs in the synthesis of amino acids

For nearly 66 years now, biotechnological processes have been utilised for the industrial synthesis of amino acids. Particularly, enzymes and whole cell biocatalysts are valuable tools for the production of proteinogenic or non-proteinogenic α-amino acids (either L- or D-enantiomer) and their derivatives which are important building blocks in the agricultural, food and pharmaceutical industries.[95] The proteinogenic and non-proteinogenic α-amino acids can be synthesised by the reductive amination of their corresponding α-keto acids catalysed by amino acid dehydrogenases (AADHs).[96] However, these processes require NADH, an expensive cofactor that is used stoichiometrically unless another enzyme is used to simultaneously regenerate it (Scheme 5-5).[96]
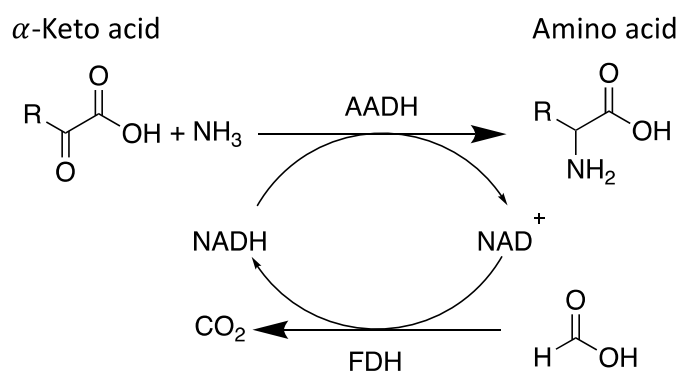
Scheme 5-5. Enzymatic synthesis of amino acids via reductive amination. AADH (amino acid dehydrogenase) and FDH (formate dehydrogenase).

For instance, L-leucine and other amino acids (such as L- isoleucine, L-alanine, L-valine and L-methionine) have been produced by reductive amination of their corresponding α-keto acids. This is catalysed by a leucine dehydrogenase (LeuDH from *Bacillus sphaericus*) in combination with a formate dehydrogenase (FDH from *Candida biodinii*) for regeneration of NADH.[97] Non-proteinogenic α-amino acids, such as L-*tert*-leucine and L-*neo*pentylglycine , have also been produced by reductive amination (Scheme 5-6).[97-99]

L-tert-leucine   L-neopentylglycine
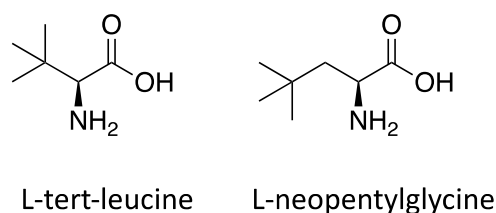
Scheme 5-6. Structures of L-tert-leucine and L-neopentylglycine.

Apart from AADHs, enzymes from the transaminase family (TAs) can be utilised for the production chiral amino containing compounds including proteinogenic and non-proteinogenic α-amino acids.[100] There are different examples in the literature showing the application of both α and ω-TAs for the synthesis of amino acids. For example, transamination of phenylpyruvate with L-aspartic acid or L-glutamic acid has been developed for the production of L-phenylalanine.[101, 102] In this route the direct transamination of α-ketoacid results in the formation of the corresponding amino acid, with no need for NADH or the addition of another enzyme for the regeneration of the cofactor (Scheme 5-7).
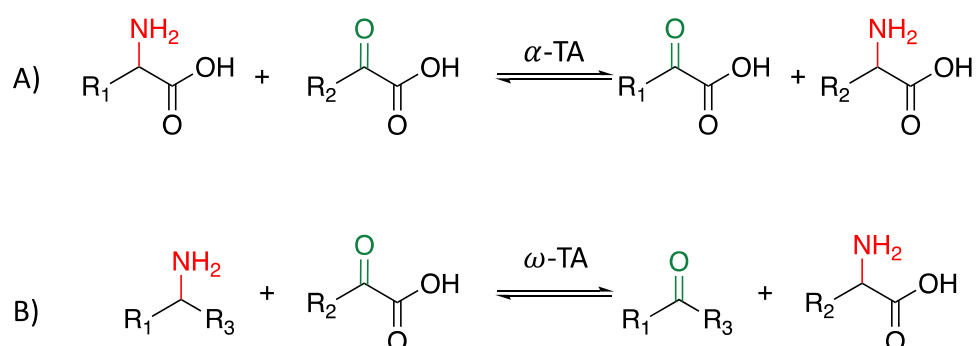
Scheme 5-7. Application of TAs in production of amino acids.

TAs are favoured biocatalysts in the production of amino acids due to the desirable features such as high turnover, relaxed substrate specificity and no need for cofactor recycling.[103] They are also powerful biocatalyst for the enantioselective synthesis of amine containing molecules

such as amine and amino acids, with high values.[104] Despite the advantages, transamination reactions are not particularly thermodynamically favourable and suffer from low conversion. This is more problematic in α-TA catalysed reactions in which the reactants and products are similar (Scheme 5-7, A). In contrast, ω-TA catalysed reactions are less thermodynamically unfavourable as the amino donor is an amine containing compound whereas the product is an amino acid (Scheme 5-7,B).[105] There are different strategies in order to overcome the thermodynamic barriers of the TA catalysed reactions. These include the use of an inexpensive amino donor in excess or removal of the reaction by-product.[94]

5.3. Results and discussion

5.3.1. Screening of commercial ATAs as potential coupling enzymes

In this thesis, a transamination step is coupled to the reaction of *Sc*PDC (Chapter 6). This second reaction is not only important to drive the equilibrium of the *Sc*PDC catalysed reaction towards carboxylation but also designed to synthesise amino acids from aldehydes and CO_2. Therefore, a transaminase enzyme is required which selectively converts the α-keto acid, the intermediate of the carboxylation reaction, and not its corresponding aldehyde. Otherwise, the concentration of the aldehyde substrate required for the carboxylation reaction is depleted which consequently affects the overall yield of coupled carboxylation-transamination reaction.
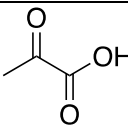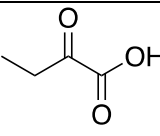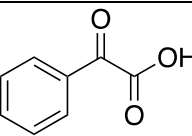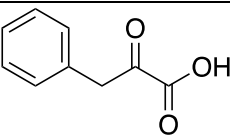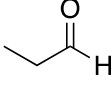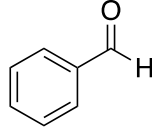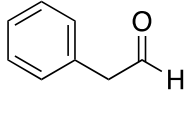
In this context, commercially available amine transaminases (ATAs, Codexis), which belong to the ω -TAs group, were initially screened with different aldehydes and their corresponding α-ketoacids. *Ortho*-xylylenediamine was employed as a 'smart' amine donor providing a fast detection method for screening of ATAs. The application of a 'smart' amine donor was firstly introduced by Green *et al.* as a novel strategy to displace the position of the unfavoured transamination reaction towards product formation.[106] Accordingly, *ortho*-xylylenediamine was applied as an amine donor in the transamination reaction, generating a by-product. This by-product undergoes an intramolecular cyclisation forming an imine followed by a spontaneous 1,5-hydride shift generating the more stable aromatic isoindole by-product. The spontaneous polymerization of the isoindole by-product forms a black precipitate which provides a high-throughput screening method (Scheme 5-8).[106]
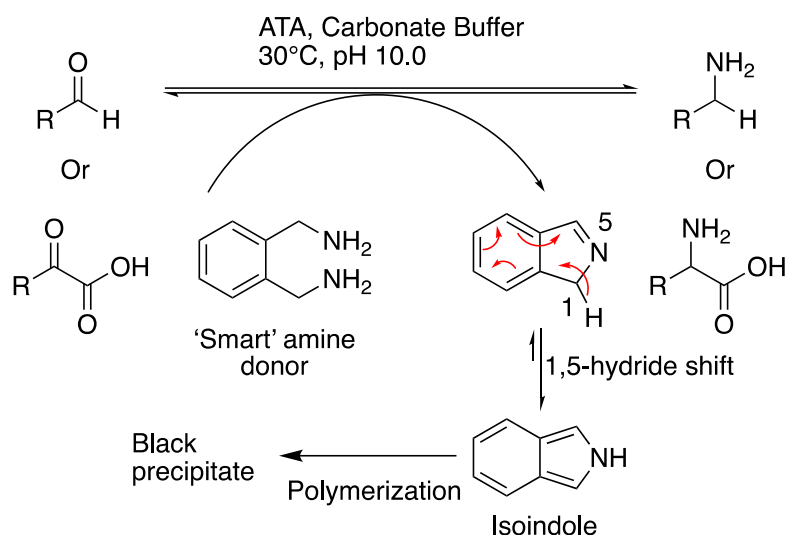
Scheme 5-8. Screening of ATAs against various aldehydes and their corresponding α-keto acids using *ortho*-xylylenediamine as a 'smart' amine donor.

Biotransformations of different aldehydes and their corresponding α-keto acids were conducted using commercially available ATAs (Codexis: ATA-113, ATA-025, ATA-0117, ATA-256) to find a transaminase enzyme that only converts the α-keto acid and not its corresponding aldehyde (Table 5-1). Aliphatic substrates including pyruvate and 2-ketobutyric acid were selected as they will be the intermediate of the carboxylation reaction catalysed by *Sc*PDC in the cascade. Also, two aromatic substrates were added to the substrate panel in order to evaluate the substrate scope of these enzymes.

Table 5-1. Chemical structure of different α-keto acids and their aldehyde used in screening with ATAs

Each reaction was set up individually in a 96-well plate by adding the amino donor (*ortho*-xylylenediamine, 10 mM) and the enzyme (0.5 mg. mL^{-1}) in carbonate buffer (90 mM, pH 10.0). The reaction was started by addition of the amine acceptor substrate (either the aldehyde or its α-keto acid, 20mM) and incubated at 30°C for 24 hours. The activity of each enzyme with different amino acceptors (aldehyde or α-keto acid) was identified qualitatively by the formation of a coloured precipitate. No colour change was an indication of no conversion. The presence of yellow colour was due to the background conversion of enzyme-bound PLP to PMP.[107]
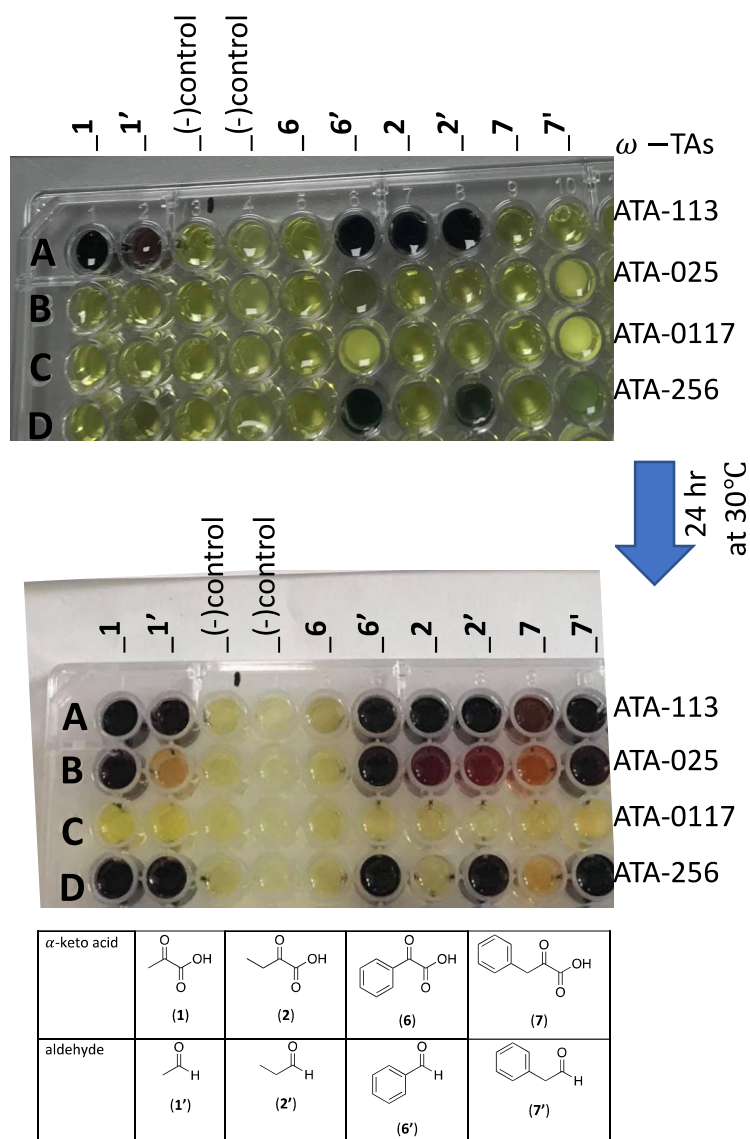
Figure 5-1. Screening of commercially available ATAs with various aldehydes and their corresponding α-keto acids in pair using *o*-xylylenediamine as the 'smart' amino donor. In the control experiments no substrate was added. The colour change in each well indicates ATA activity.

As shown in Figure 5-1, with ATA-0117 all wells remained yellow indicating that this enzyme was unable to accept any of the aldehydes or their corresponding α-keto acid as the substrate (C1-C4'). Also, all of the ATAs applied were inactive towards phenylglyoxylic acid (A2-D2). In the control experiments where no amino acceptor was added the wells remained yellow as expected (A-D (-) control). In some wells the formation of black precipitate was observed immediately after the addition of the amino acceptor substrates (A1, A1', A2', A3, A3', D2' and D3') indicating that these reactions proceed at higher rate. After 24hr, significant colour changes developed in most wells, showing that the biotransformation proceeded with high conversion in these wells. However, in some other wells, less intense coloration was observed showing a moderate level of conversion (B1', B3, B3', B4 and D4). For example, ATA-025 with acetaldehyde turned to orange after 24 hours showing that the enzyme is not selective with pyruvate compared to acetaldehyde. Among all the ω-TAs tested, we couldn't identify an enzyme that selectively converts the α-keto acid and not its corresponding aldehyde. Consequently, the idea of using commercially available ATAs in the coupled carboxylation-transamination reaction was discarded.

5.3.2. *Sc*ALT as potential coupling enzyme: Cloning

Alanine aminotransferase (ALT; EC2.6.1.2, previously named glutamate pyruvate transaminase [GPT]) is a PLP dependent enzyme catalysing the reversible transfer of an amino group from alanine to 2-oxoglutarate to form pyruvate and glutamate (Scheme 5-9).[108]
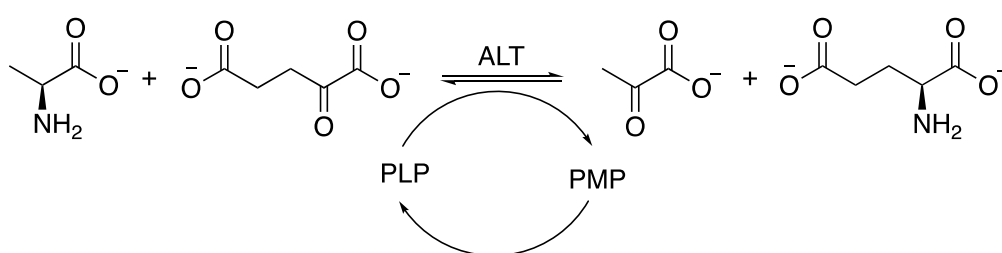
Scheme 5-9. Transamination reaction catalysed by ALT

In the yeast *S. cerevisiae*, there are two genes that encode for alanine aminotransferase, *Sc*ALT1 and *Sc*ALT2.[109] Despite having 65% identity in their protein sequences , they are very different proteins in terms of their function. *Sc*ALT1 is known to be involved in the biosynthesis of alanine, whereas *Sc*ALT2 does not have any aminotransferase activity.[110] These differences have been attributed to their interactions with the PLP within the active

site. Tertiary structural analysis of both enzymes has revealed that *Sc*ALT2 has a more open conformation, which subsequently affects the polarity of the active site, causing each enzyme to have a different interaction with PLP.[111] *Sc*ALT1 is known to accept pyruvate as the substrate and catalyse the transfer of amino group from glutamate to pyruvate (K_m =0.4 mM and k_{cat}=8.64 s^{-1}) to form L-alanine.[51] Therefore, *Sc*ALT1 was evaluated as the coupling enzyme later in combination with *Sc*PDC in a one-pot carboxylation-transamination reaction for the synthesis of L-alanine. 2-ketobutyric acid is the closest structure to pyruvate and it could serve as a keto substrate for *Sc*ALT1. Here, the transaminase activity of *Sc*ALT1 with 2-keto butyric acid was also assessed for the synthesis of L-Homoalanine.

For this purpose, *Sc*ALT1 was cloned using the *S. cerevisiae* genomic DNA as prepared before (Section 2.1). The *Sc*ALT1 gene was amplified from genomic DNA of *S. cerevisiae* and inserted into the pJexpress414 expression vector using NEBuilder HiFi DNA assembly to create the *Sc*ALT1-pJexpress414 plasmid. According to the NEBuilder HiFi DNA assembly Cloning kit (NEB) guideline, the vector and the insert may be viewed as two PCR fragments that have to be assembled into a circular DNA molecule. Therefore, the primers were designed in such a way to share overlapping ends. To amplify the *Sc*ALT1 gene, both forward and reverse *Sc*ALT1-specific priming sequences (grey) at their 5' end were fused with the respective pJexpress414 sequences to be used as overlap sequences in assembly with the vector. Within the *Sc*ALT1 forward PCR primer, the overlap sequence (orange) was identical to the 20-nt terminal sequence on the top strand of pJ414 left arm (in the 5' to 3' direction) (thrombin site). In the *Sc*ALT1 reverse PCR primer, the overlap sequence (blue) was identical to the 20-nt terminal sequence on the bottom strand of the pJexpress414 right arm (Figure 5-2).
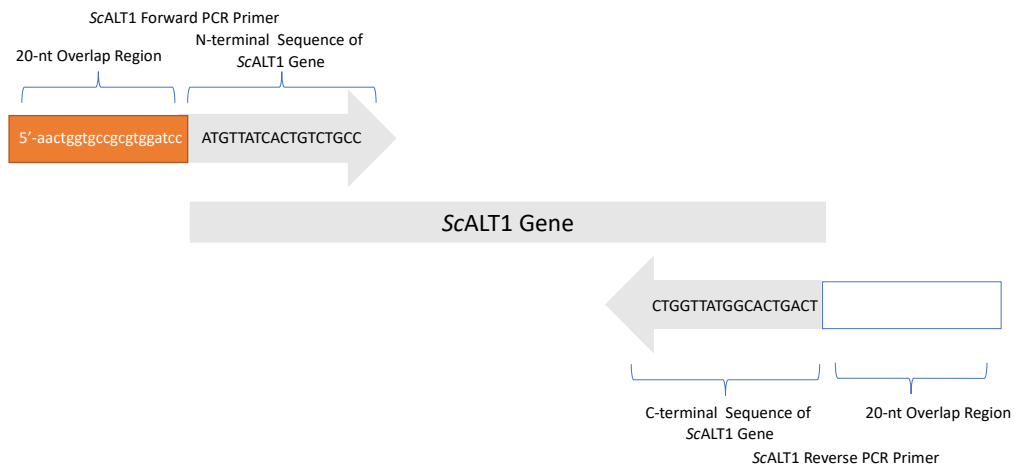
Figure 5-2. Primer design for amplification of *Sc*ALT1 gene and incorporation of overlapping regions

The PCR reaction for amplification of the *Sc*ALT1 fragment was followed using the *Sc*ALT1 forward and reverse PCR primers at the Ta (annealing temperature) of 59.5 °C as recommended by the manufacturer. Two individual PCR reactions were set up by using primers at the final concentrations of 0.5 and 0.25 μM.
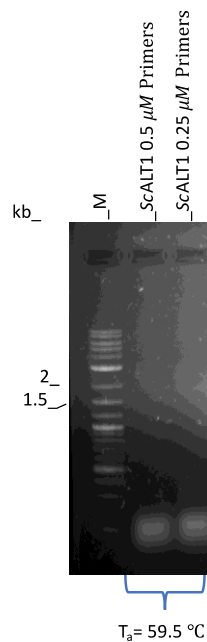
Figure 5-3. Amplification of *Sc*ALT1 fragment from *S. Cerevisiae* by PCR analysed by 1.2% (w/v) agarose gel. M represents DNA ladder (1Kb Plus DNA ladder, N0550G)

Following each PCR reaction, the amplification of *Sc*ALT1 fragment (1.8 kb) failed to generate any PCR product (Figure 5-3). Due to the inadequate amounts of DNA being obtained using the manufacturer's protocol several modifications were made to the protocol. These

variations included changes in the concentrations of the primers and annealing temperature. The concentration of primers was varied between 0.3 μM and 0.15 μM. Also, the different annealing temperatures (Ta= 55, 61, 58.5, 72 $°C$) were applied.
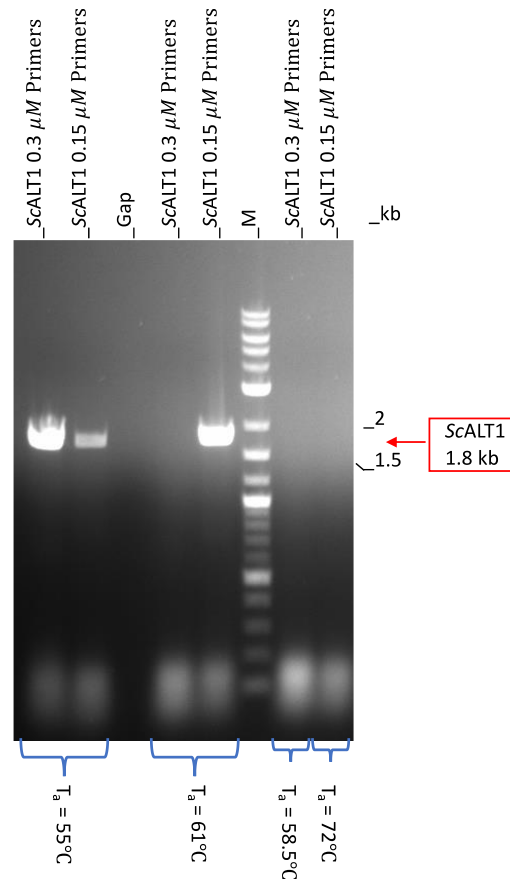
Figure 5-4. Analysis of PCR amplification of *Sc*ALT1 fragment from *S. Cerevisiae* by 1.2% (w/v) agarose gel using different primer concentrations and annealing temperatures. M represents DNA ladder (1Kb Plus DNA ladder, N0550G)

A strong amplification band for *Sc*ALT1 was observed when primers used at the final concentrations of 0.3 μM at 55 $°C$ annealing temperature (Figure 5-4, lane 1 from left). Also, amplification with 0.15 μM primers at annealing temperature of 61 $°C$ showed a strong band at the correct size (Figure 5-4, lane 5 from left). The PCR product obtained using the condition in the lane 1 was taken for the assembly of *Sc*ALT1 into the pJ414 backbone.

The pJ414 backbone was linearized using the pJ414-*Sc*PDC vector as the template. The primers (pJ414-backbone FWD and pJ414-backbone REV) were designed in such a way to exclude *Sc*PDC fragment from the vector. The pJ414 vector with the *Sc*PDC insert has the size of 5.8 kb in total. Upon removing the *Sc*PDC fragment (1.7 kb), the size of 4.1 kb was expected

for the linearised vector. Analysis of the PCR reaction showed the expected pJ414 linearized back bone (Figure 5-5).
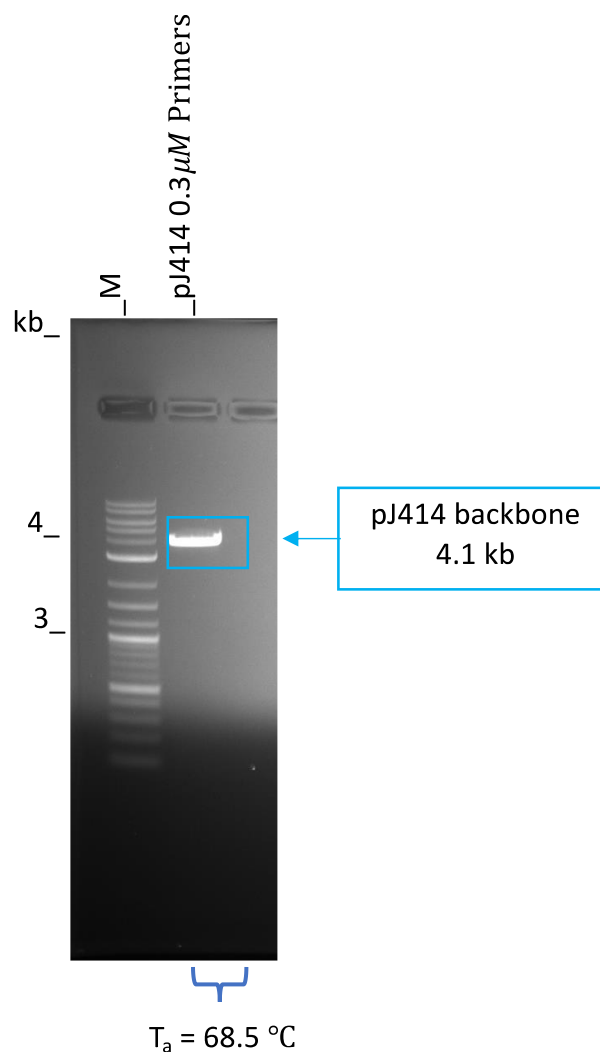
Figure 5-5. Gel Electrophoretic analysis of pJ414 linearization by 1.2% (w/v) agarose gel . M represents DNA ladder (1Kb Plus DNA ladder, N0550G).

ScALT1 fragment and pJ414 linearized backbone were assembled using the NEBuilder HiFi DNA assembly kit. The two fragments were mixed in 3:1 ratio (ScALT1: pJ414) in DNA Assembly Master Mix buffer which includes different enzymes: exonucleases, DNA polymerase and ligase. The mix was incubated at 50 °C at three different incubation times (15 minutes, 1 and 4 hours). The control (negative) reactions were performed by excluding the master mix from the reactions. The analysis of each reaction showed the successful assembly of ScALT1 and pJ4141 fragments (5.9 kb) (Figure 5-6).
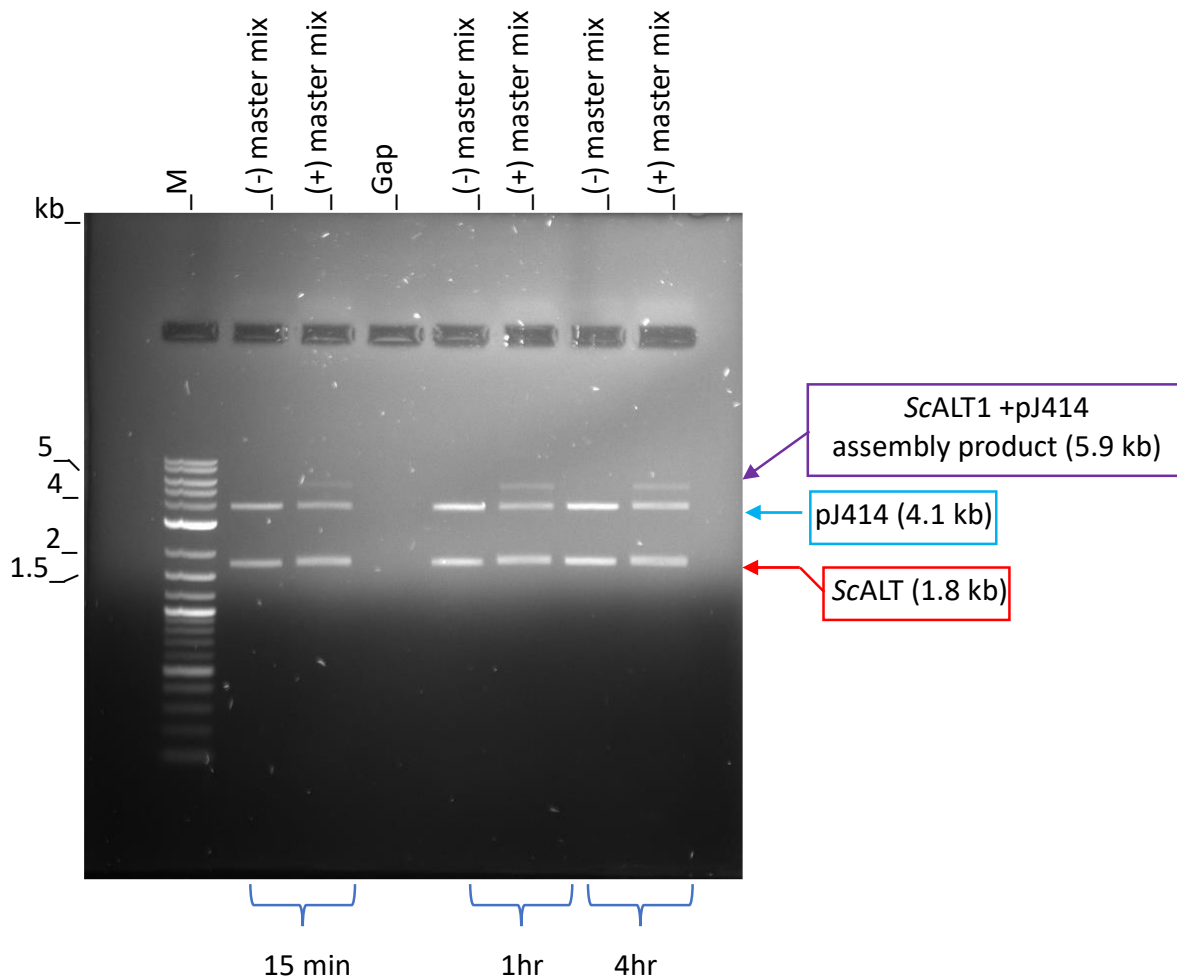
Figure 5-6. Assessment of *Sc*ALT1 and pJ414 assembly by 1.2% (w/v) agarose gel using NEBuilder HiFi DNA assembly kit. M represents DNA ladder (1Kb Plus DNA ladder, N0550G).

Following the assembly, the circular DNA (*Sc*ALT1-pJ414 vector) was transformed into NEB storage strain (NEB5alpha) and plated on an ampicillin plate. The transformant colonies grown on the ampicillin plate were picked and screened by PCR using *Sc*ALT1 forward and reverse primers as such colonies were considered to likely harbour a pJexpress414 derived plasmid with a *Sc*ALT1 gene correctly inserted. These primers anneal at temperature of 72 °C to the *Sc*ALT1-pJ414 construct. After PCR screening all products were examined by 1.2% (w/v) agarose gel (Figure 5-7). The PCR screening results suggested that the ScALT1 gene was correctly inserted into pJ414express backbone. Finally, the DNA extracted from colony 4 with T7 forward and pJexpress414 reverse primers were submitted for sequencing. The sequencing results confirmed the replacement of *Sc*PDC gene by a 1.8 kb *Sc*ALT1 gene. The original T7 promoter, His tag located at the N-terminal remained in the plasmid.
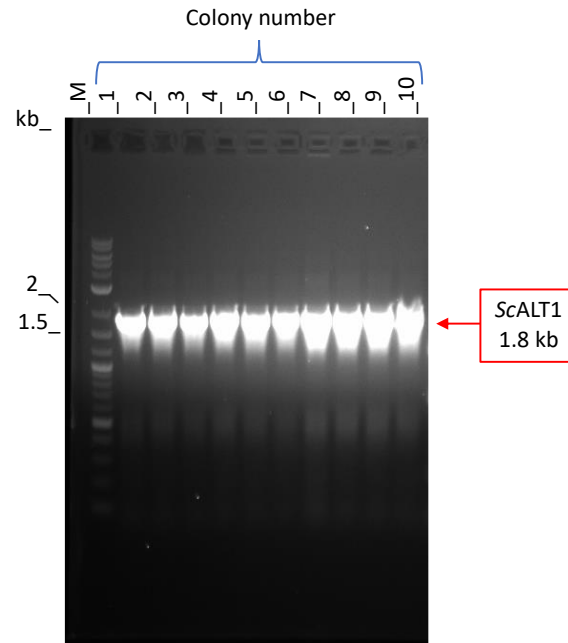
Figure 5-7. PCR Screening of colonies harbouring *Sc*ALT1-pJ414 plasmid. M represents DNA ladder (1Kb Plus DNA ladder, N0550G).

5.3.3. Recombinant over-expression of *Sc*ALT1 in *E. coli*

The over-expression of *Sc*ALT1 was followed according to the procedure reported by Duff *et al.,* using BL21(DE3) *E. coli* host cells and IPTG at the final concentration of 1mM in LB media.[51] *Sc*ALT1-pJ414 expression construct was used to transform the BL21 *E. coli* expression host. Initial expression tests were performed on a small scale (20 ml) in order to evaluate the extent of induction and solubility of *Sc*ALT1 when expressed as a recombinant protein in *E. coli.* Recombinant *Sc*ALT1 was expressed as a His tag (N-terminal) protein following the addition of 1mM IPTG. Different temperatures and incubation times were tested as these parameters weren't described in the method reported by Duff *et al.*[51] These included the growth of cells at 37 °C for 4 hours, 37 °C for 18 hours, 30 °C for 18 hours and 20 °C for 18 hours. Following induction, bacterial cells were harvested and lysed chemically. Soluble and insoluble fractions were separated. All fractions including non-induced, induced soluble and induced insoluble were analysed by 12% (w/v) SDS-PAGE gel electrophoresis (Figure 5-8). However, the *Sc*ALT1 protein band with the expected molecular weight of 69.7 kDa wasn't detected in any of the conditions.
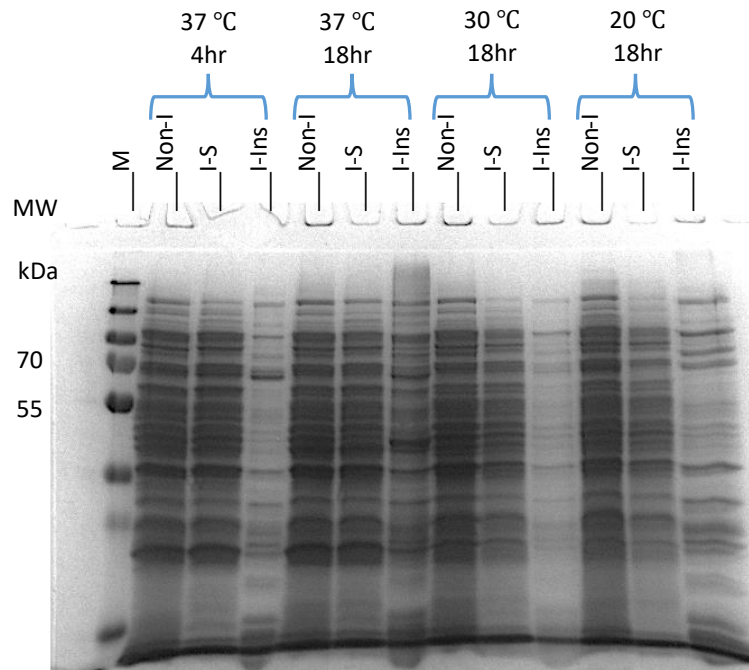
Figure 5-8. SDS-PAGE analysis of recombinant expression of *Sc*ALT1 in *E. coli* BL21 cells. Total protein was extracted from *E. coli* cells with no IPTG added (Non-induced: Non-I), with IPTG added (Induced) (1mM) at various temperatures and incubation times. I-S (induced-soluble), I-Ins (Inducted-insoluble) and M (Marker: PageRuler™ Plus Prestained Protein Ladder from ThermoFisher).

Since the over-expression in BL21 cells failed to generate *Sc*ALT1 proteins, bacterial host strain was changed from BL21 to Rosetta *E. coli* cells. This strain has been particularly designed to enhance the expression of proteins that have high frequencies of rare codon. The sequence analysis of *Sc*ALT1 revealed that there are rare codons with high frequency of occurrence including: AGG (7), AGA (11), AUA (9), CUA (12), CCC (4), GGA (2). The enhancement in the expression is facilitated by the presence of a plasmid within the Rosetta *E. coli* cells that encodes rare tRNAs including, AGG, AGA, AUA,CUA,CCC and GGA.[112] Also, over-expression of *Sc*ALT1 in Rosetta *E. coli* cells was previously reported.[51] The protein expression was conducted at 37 °C and monitored at 1hr, 2hr, 3hr and 18 hr using the same concentration of IPTG (1mM) as before. All fractions including non-induced, induced soluble and induced insoluble were analysed by 12% (w/v) SDS-PAGE gel electrophoresis (Figure 5-9). *Sc*ALT1 protein band (69.7 KDa) was detected showing the development of protein expression as the time of incubation increased. However, the protein was mainly in the insoluble fraction.
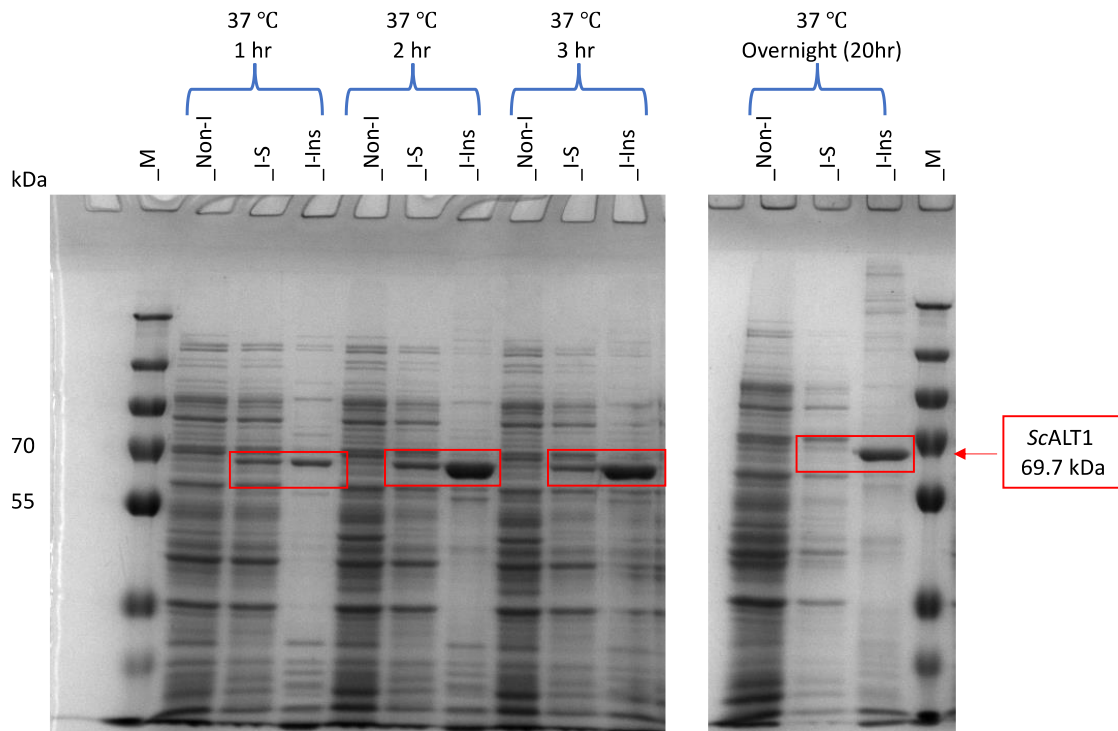
Figure 5-9. SDS-PAGE analysis of recombinant expression of *Sc*ALT1 in *E. coli* Rosetta cells. Total protein was extracted from E. coli cells with no IPTG added (Non-induced), with IPTG added (Induced) (1mM) at 37 °C and various incubation time. I-S (induced-soluble), I-Ins (Inducted-insoluble) and M (Marker: PageRuler™ Plus Prestained Protein Ladder from ThermoFisher).

To maximise the efficient production of soluble *Sc*ALT1 in Rosetta *E. coli* cells, the concentration of IPTG was lowered from 1 mM to 0.1 mM. However, the insoluble expression was still observable even using lower concentration of the IPTG (Figure 5-10).
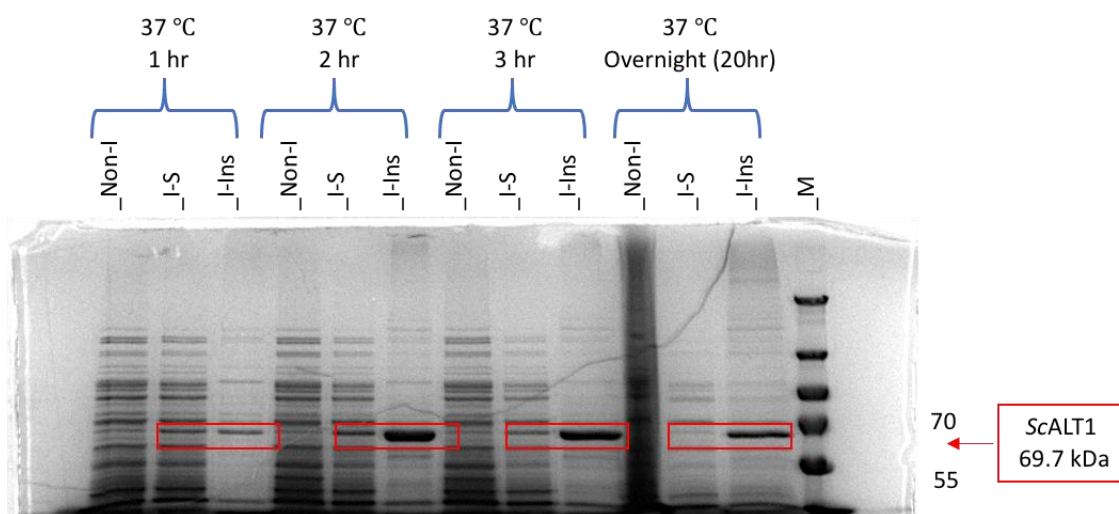
Figure 5-10. SDS-PAGE analysis of recombinant cxpression of *Sc*ALT1 in *E. coli* Rosetta cells. Total protein was extracted from *E. coli* cells with no IPTG added (Non-induced), with IPTG added (Induced) (0.1 mM) monitored at 37 °C.

In another experiment, both IPTG concentration and the temperature were lowered. The final concentration of IPTG was 0.1 mM and the cells were grown at 15 °C over night. Also, expression in M9 media was attempted under the exact same conditions. Only limited soluble expression was observed using 0.1 mM IPTG at 15 °C in the LB media. Lower level of expression was observed in minimal media (Figure 5-11).
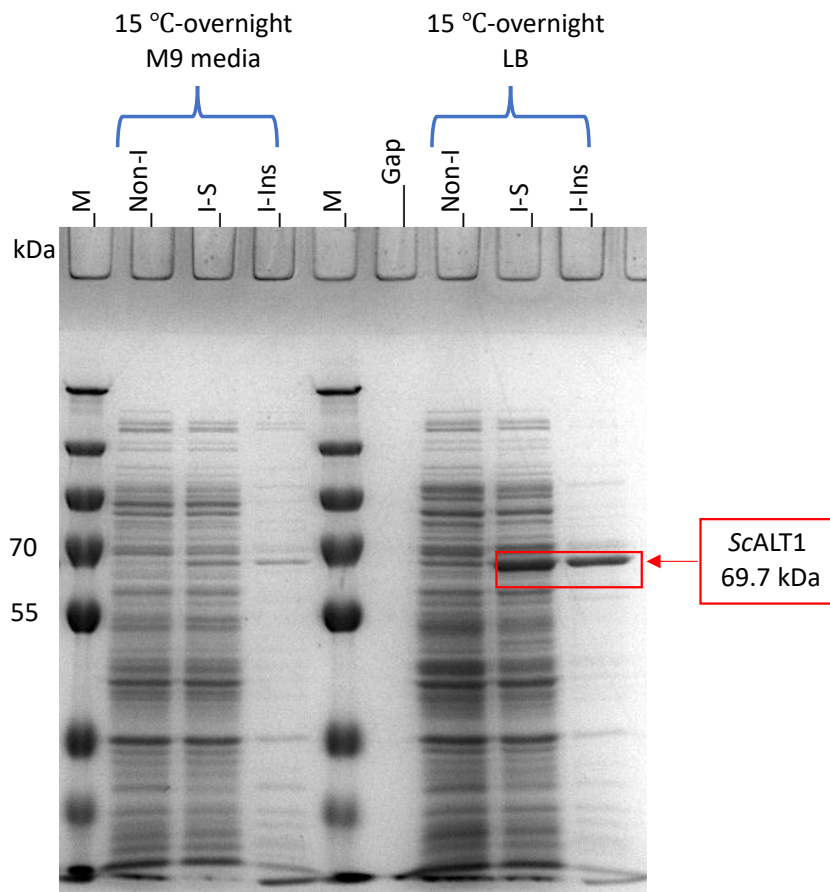
Figure 5-11. SDS-PAGE analysis of recombinant expression of *Sc*ALT1 in *E. coli* Rosetta cells. Total protein was extracted from *E. coli* cells with no IPTG added (Non-induced), with IPTG added (Induced) (0.1 mM) at 15 $°C$ in both M9 media and LB. I-S (induced-soluble), I-Ins (Inducted-insoluble) and M (Marker: PageRuler™ Plus Prestained Protein Ladder from ThermoFisher).

In the published report by Ortega *et al.*, the mitochondrial localization sequence was excluded from the *Sc*ALT1 protein sequence .[111] However, no explanation was provided for such alteration. It is well known that many of the proteins that are targeted to mitochondria have a mitochondrial localization sequence which is usually present at the N-terminus of the protein.[113] Once the protein is targeted to mitochondria the signal peptide (MSP) is removed from the mature protein by the action of peptidase. One of the strategies to improve the solubility of a protein is to remove a part of sequence, which is not involved in the catalytic

activity of the protein, and this can be a MSP.[114] Here the effect of removal of the signal peptide on the solubility of the *Sc*ALT1 was investigated when overexpressed in Rosetta *E. coli* cells and compared with the previous results. The removal of mitochondrial localization sequence from the *Sc*ALT1 sequence was performed by site directed mutagenesis. The mutagenic primers were designed to remove 240 base pairs (80 amino acids) at N-terminus of the sequence corresponding to the mitochondrial localization sequence (Figure 5-12). The deletion was generated by designing forward and reverse primers that flank the MSP site to be deleted. Removal of the sequence was verified by sequencing.
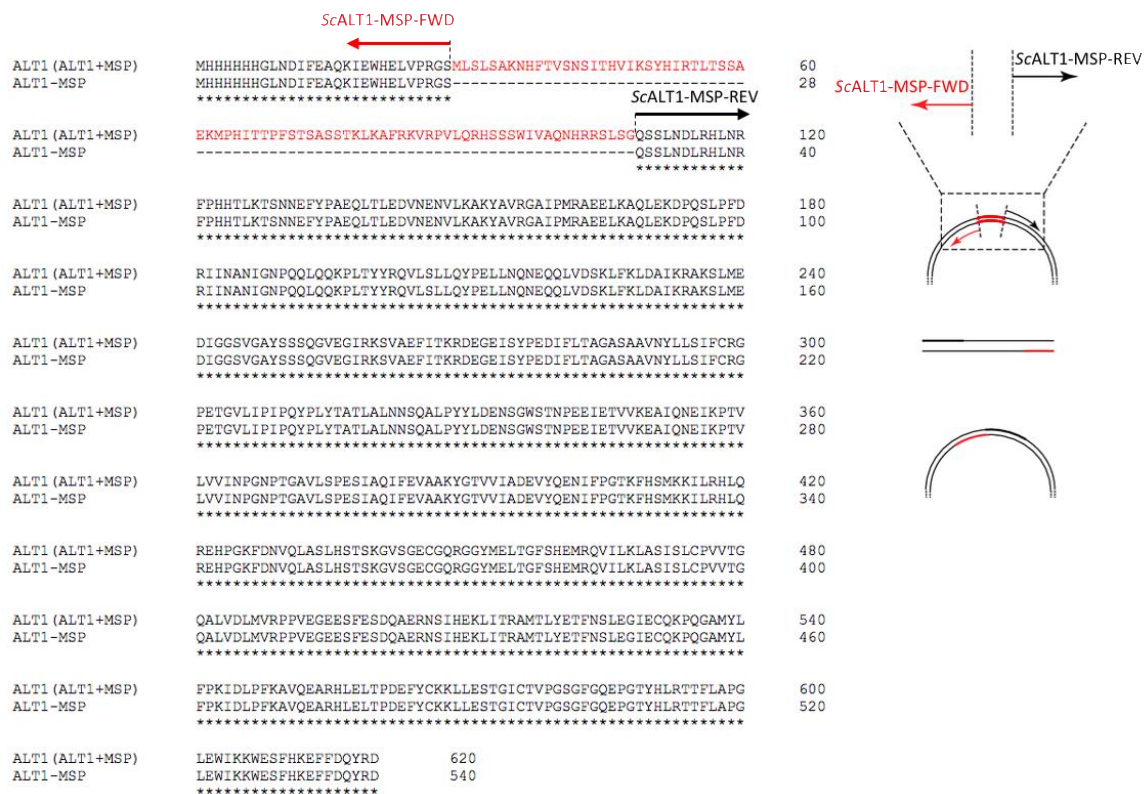
Figure 5-12. Sequence alignment of ALT with (ALT+MSP) and without (ALT-MSP) Mitochondrial signal peptide

The new construct was then transfected into Rosetta *E. coli* strain. In order to assess the effect of removing the mitochondrial localization sequence on the expression level and solubility of the protein, initial expression tests for both *Sc*ALT1 with the mitochondrial localization sequence (*Sc*ALT1+MSP) and without the mitochondrial localization sequence (*Sc*ALT1-MSP) were conducted in parallel on a small scale (20 ml). Recombinant proteins were expressed in LB medium following the addition of 100 μM IPTG and the growth of cells at 15 °C overnight. Following induction, bacterial cells were harvested and lysed chemically. For *Sc*ALT1+MSP,

bands of the expected size (69.7 kDa) were observed in both insoluble and soluble fractions following induction. For *Sc*ALT1-MSP following the induction, band of the expected size (60.8 KDa) was only detected in the soluble fraction (Figure 5-13). Also, the shift in the molecular mass of protein confirmed the removal of the mitochondrial localization sequence. Therefore, the expression of the *Sc*ALT1 as a soluble protein was substantially improved after excluding the mitochondrial signal peptide.
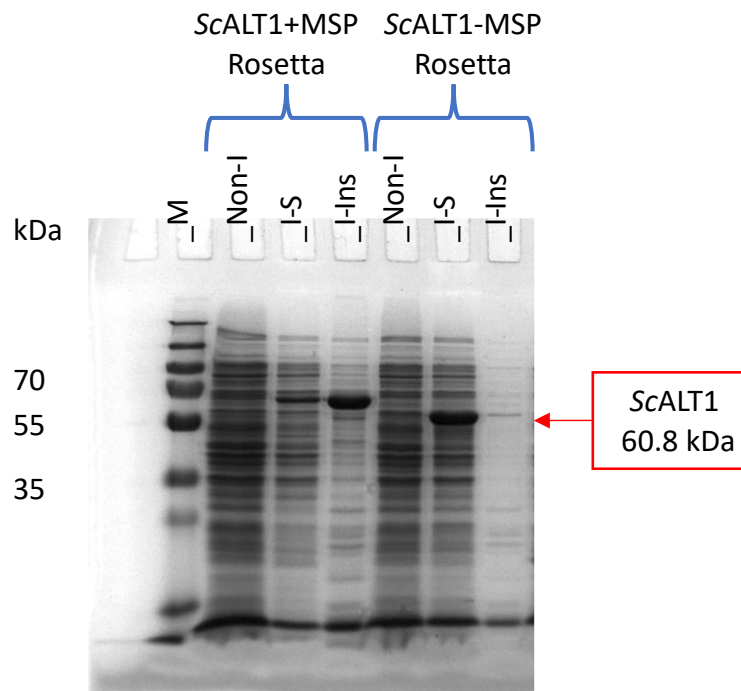
Figure 5-13. SDS-PAGE analysis of recombinant expression of ScALT1 with (*Sc*ALT1+MSP) and without (*Sc*ALT1-MSP) mitochondrial signal peptide in Rosetta *E. coli* cells. Total protein was extracted from E. coli cells with no IPTG added (Non-induced), with IPTG added (Induced) (100 μM) at 15 $°C$. I-S (induced-soluble), I-Ins (Induced-insoluble) and M (Marker: PageRuler™ Plus Prestained Protein Ladder from ThermoFisher).

5.3.4. Purification of recombinant *Sc*ALT1

After the optimisation of the conditions for expression of *Sc*ALT1-MSP, the expression was scaled-up to a 1 litre culture. The Rosetta *E. coli* cells were transformed with *Sc*ALT1-MSP plasmid. Recombinant proteins were expressed in LB medium following the addition of 0.1 mM IPTG and the growth of cells at 15 °C overnight. Cell pellets containing *Sc*ALT1-MSP were centrifuged and stored in − 80 °C until needed for purification. Bacterial cell pellets were re-suspended in buffer D and sonicated. *Sc*ALT1-MSP was purified using IMAC. *Sc*ALT1-MSP bound to the column was washed with 3 %, 5%, 8%, 30% buffer E containing. (His)$_6$-*Sc*ALT

bound to the column was eluted with 50% buffer E. Main fractions containing *Sc*ALT1-MSP were identified by SDS-PAGE and combined (Figure 5-14). The combined fractions were concentrated, and buffer exchanged to buffer F containing using vivaspin 20 centrifuge column (30 kDa MWCO). The sample was aliquoted and stored in – 80 °C for future use.
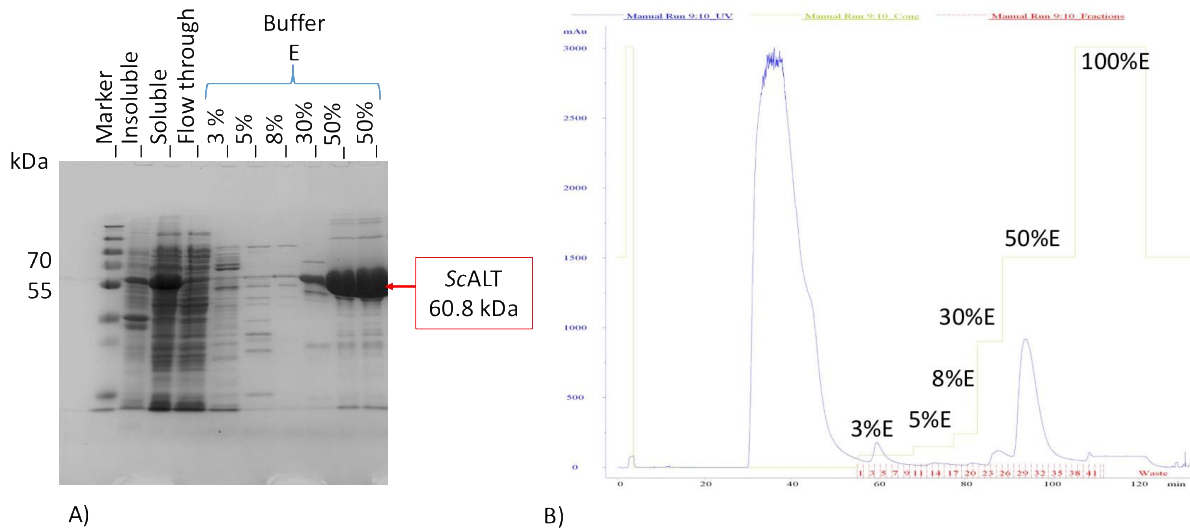
Figure 5-14. Expression and purification of (His)$_6$-*Sc*ALT protein using Ni^{2+} affinity chromatography and AKTA technology. (A) Comassie stained SDS-PAGE gel (12%) of the proteins present in different fractions including insoluble, soluble, flow-through, wash steps and finally elution with different concentrations of imidazole, Marker (PageRuler Prestained Protein Ladder) (B) Representative chromatogram for purification of (His)$_6$-*Sc*ALT recombinant protein using Ni^{2+} affinity chromatography and AKTA technology. Absorbance at 280 nm is shown in blue and the concentration of the buffer B used in green.

5.3.5. Transamination of α-keto acids with *Sc*ALT1

The aminotransferase activity of purified *Sc*ALT1 was measured by HPLC assay. This was performed by measuring the amount of the expected amine product using a calibration curve of authentic standards (L-alanine and L- homoalanine 0-25 mM derivatised with OPA, section 2.14.2.) and calculating the yield.
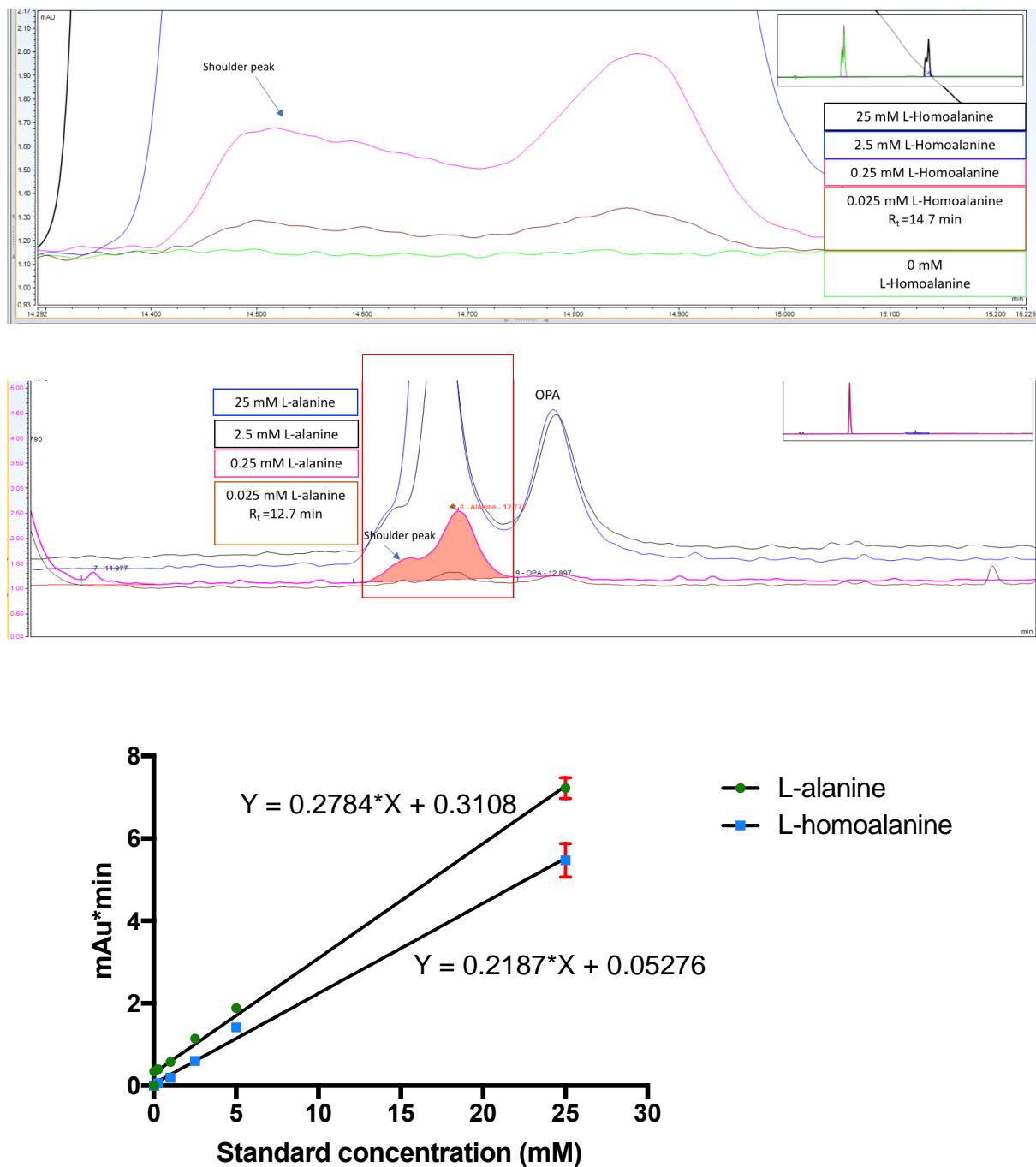
Figure 5-15. HPLC Chromatograms and calibration curve of L-alanine and homoalanine derivatised with OPA. For each standard the integration of the entire peak including the shoulder was used to build the calibration curve. Error bars in red indicate standard deviations (SD) for three repeat experiments.

The transamination of two different α-keto acid (pyruvate and 2-ketobutyric acid) catalysed by *Sc*ALT1, was tested independently in the carbonate buffer (250 mM, pH 8.5) by incubating the enzyme with the amino donor (glutamic acid, 25 mM) and the amino acceptor (pyruvate or 2-ketobutyric acid, 2.5 mM). After the incubation period, formation of the desired amino

acids (L-alanine and L-homoalanine) was observed, and this confirmed the ability of *Sc*ALT1 to efficiently produce L-alanine and L-homoalanine from their respective α-keto acids (Figure 5-16).

The higher yield was observed with pyruvate being the natural substrate of the enzyme. There are no reported values on the *Sc*ALT1 activity with 2-ketobutyric acid. The only example that has turned over this substrate is the ALT from *Candida maltosa* (*C. maltosa* ALT) with only 32.4% relative activity compared to pyruvate .[115] The reaction was also analysed after 48 hours with no significant changes in the yield of the amino acid product indicating that the reaction has reached equilibrium.
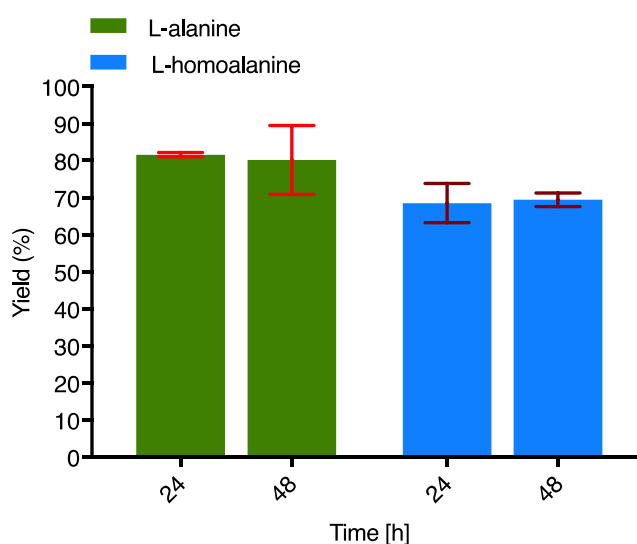
Figure 5-16. Transamination step for conversion of 2-ketoacids (pyruvate and 2-ketobutyric acid) catalysed by *Sc*ALT1 after 24 and 48 hours in 250 mM carbonate buffer pH 8.5. The peak appearance of each product, L-alanine and L-homoalanine, was similar to the standards showing the shoulder peak. The entire peaks including the shoulder were integrated. Error bars in red indicate standard deviations (SD) for three repeat experiments.

5.3.6. Effect of pH on ScALT1 transaminase reaction

The effect of pH on the transamination reaction catalysed by *Sc*ALT1 was examined. Each reaction was performed separately in two different buffers having different pH's ranges: 250 mM sodium phosphate (NaH$_2$PO$_4$-Na$_2$HPO$_4$, pH 6.0-7.5) and 250 mM sodium carbonate (NaHCO$_3$-Na$_2$CO$_3$, pH 8.5-10.5) using 2-ketobutryic acid as the amino acceptor and glutamate as the amino donor. Each reaction was initiated by addition of the *Sc*ALT1. The yield of the amino acid product was measured by HPLC. As illustrated in Figure 5-17, the enzyme showed

slightly higher activity in the alkaline pH's (8.5 - 9.5), with the maximum activity at pH 9.0. No transaminase activity was observed at pH 10.0.
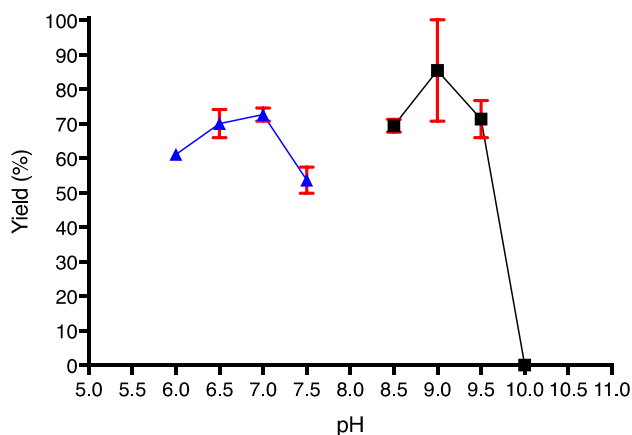
Figure 5-17. Effect of pH on the yield of the transamination step of 2-ketobutyric acid catalysed by *Sc*ALT after 48 hours. Triangle (250 mM sodium phosphate) and square (250 mM carbonate) buffer. Error bars in red indicate standard deviations (SD) for three repeat experiments.

5.3.7. Effect of aldehyde, the substrate of carboxylation reaction, on ScALT activity

Both acetaldehyde and propanal could be substrates for *Sc*PDC catalysed carboxylation reaction. Therefore, it was important to check the effect of these aldehydes on the ScALT transaminase activity before attempting the coupled reaction. Two separate reactions were set up separately as described in (section 5.3.5) in the presence of each aldehyde (acetaldehyde or propanal, 25 mM). After 24 hours, each reaction was analysed by HPLC showed no changes in *Sc*ALT activity. Therefore, the transfer of the amine group from glutamate to the amine acceptor (pyruvate or 2-ketobutryic acid) was not inhibited in the presence of the corresponding aldehyde (acetaldehyde or propanal).

5.4. Conclusion

In this chapter, enzymes from the transaminase family were selected as potential coupling enzymes to be used later to in combination with *Sc*PDC in a one-pot reaction for the synthesis of amino acids from aldehydes and CO_2. In our initial screening, commercially available ATAs were not fit for our purpose due to their lack of selectivity. An α-TA from *S. cerevisiae* was selected as our next candidate. This selection was based on the known activity of this enzyme in the natural biosynthesis of amino acids, particularly L-alanine, in this organism. *Sc*ALT1 gene

was successfully amplified and cloned into pJexpress414 vector. The expression of the enzyme in *E. coli* cells was tested and optimised. The *Sc*ALT1 was purified as a His-tag protein and it was used to transaminate different substrates with a very high yield. Here, the aminotransferase activity of *Sc*ALT1 with 2-ketobutryic acid was tested for the first time, showing that the enzyme accepts this substrate as well as pyruvate. To drive the equilibrium of the transamination reaction towards production formation 10:1 ratio of amino donor to amino acceptor were used. The results showed that the reaction reaches equilibrium after 48 hours. This was crucially important for our one-pot carboxylation-transamination reaction in order to drive the equilibrium of the reaction catalysed by *Sc*PDC towards carboxylation. The activity of the *Sc*ALT was tested in different buffers with preferred activity in alkaline conditions. This would be beneficial later as the carboxylase activity of the *Sc*PDC has also been reported in alkaline conditions.[45] Also, the yield of *Sc*ALT1 transamination reaction was not influenced by the presence of aldehydes, the substrate for the carboxylation reaction catalysed by *Sc*PDC.

Chapter 6. Biocatalytic synthesis of the amino acid L-homoalanine by exploiting the CO₂ fixation of an α-keto decarboxylase

6.1. Background

As discussed earlier (Chapter 3 and Chapter 4) the carboxylation reactions of two decarboxylase enzymes (*Sc*PDC and *Ll*kdcA) could not be achieved in a single step. Here a one-pot carboxylation-transamination reaction is designed not only to drive the unfavoured carboxylation reactions but also to convert CO_2 into valuable chemicals. Our work has focused on a novel biocatalytic approach in which L-homoalanine is synthesised from CO2 and propanal using *Sc*PDC to catalyse the formation of 2-ketobutyric acid, which is then converted to L-homoalanine in a coupled reaction with *Sc*ALT. Detailed studies of this biocatalytic approach have provided an in-depth understanding of the parameters which directly influence the yield of the reaction.

L-Homoalanine is an important non-proteinogenic amino acid which has been used as a key precursor for the synthesis of enantiopure pharmaceutical agents, including levetiracetam (trade name Keppra) and brivaracetam, two antiepileptic medicines, and ethambutol, a compound used for the treatment of tuberculosis (Scheme 6-1).[116-118]
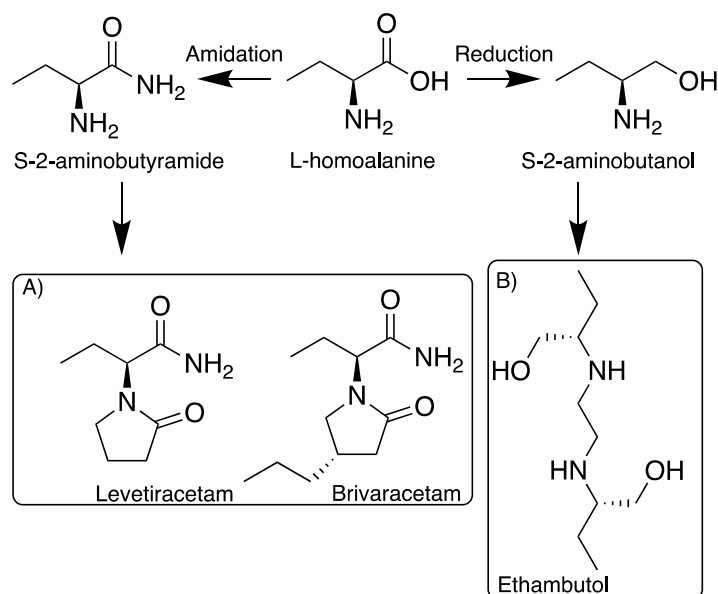
Scheme 6-1. Utilisation of L-homoalanine in pharmaceutical industry. The chemical synthetic routes for the production of antiepileptic (A) and anti-tuberculosis (B) medicine(s) from L-homoalanine.

Biocatalytic synthesis of the amino acid L-homoalanine by exploiting the CO2 fixation of an α-keto decarboxylase

L-Homoalanine can be synthesised chemically from propanal via the Strecker reaction.[119] However, the chemical synthesis route to L-homoalanine requires harsh reaction conditions (high temperature 90 °C), organic solvents such as chloroform, as well as a highly toxic starting material, sodium cyanide (Scheme 6-2). These make the process both hazardous and non-environmentally friendly.[119] In addition, the stereospecificity of the Strecker reaction can only be directed by a chiral auxiliary, such as [(1S)-1-(4-methoxyphenyl)ethyl] amine hydrochloride, which significantly decreases the atom efficiency of the reaction.[119, 120] The separation of the reaction intermediate, 2-[1-(4-Methoxyphenyl)-1-(S)-methyl-ethylamino]-(S)-butyronitrile hydrochloride, followed by an extraction step to obtain the final product as well as not recycling the chiral auxiliary are against the 12 principles of green chemistry.[121]
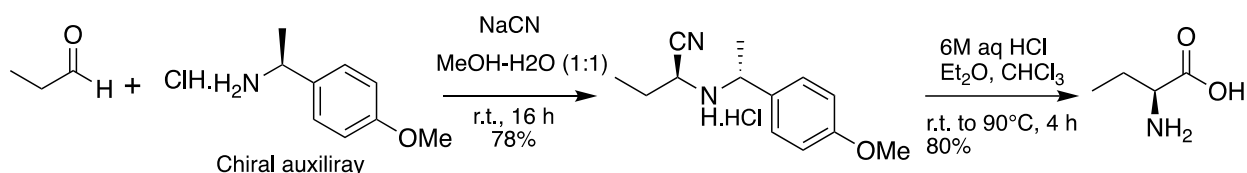
Scheme 6-2. Chemical synthesis of L-homoalanine via the Strecker reaction.[119]

Therefore, the development of a green biocatalytic approach is highly desirable, not only because it is more sustainable, but also fixes CO_2. Our approach consists of a one-pot enzymatic cascade involving two steps: carboxylation and transamination (Scheme 6-3). In the first step 2-ketobutyric acid is synthesised from CO_2 and propanal using the reverse of the natural reaction catalysed by ScPDC followed by a transamination step catalysed by *S. cerevisiae* L-alanine transaminase (ScALT). The second step of the cascade is necessary not only to generate the desired product but also to shift the position of the equilibrium of the first step towards the unfavoured carboxylation, as determined by Le Chatelier's principle.
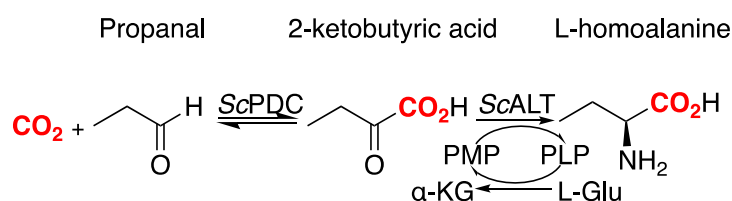
Scheme 6-3. One-pot two step enzymatic synthesis of L-homoalanine from CO_2 and propanal. The product of the carboxylation step, 2-ketobutryic acid, is immediately converted to L-homoalanine using L-Glu as the amine donor.

6.2. Results and discussion:

6.2.1. Effect of CO$_2$ pressure on the yield of L-homoalanine synthesis

In order to investigate whether the carbonate buffer is an accessible source of CO$_2$ to achieve CO$_2$ fixation without the need for a gaseous source of CO$_2$, the proposed enzymatic route was examined under atmospheric pressure. Therefore, in our initial experiments, the synthesis of L-alanine and L-homoalanine were attempted under atmospheric pressure using carbonate buffer (NaHCO$_3$-Na$_2$CO$_3$, 250 mM, pH 8.5) as the source of CO$_2$. After 48 hours, neither of the amino acids, namely L-alanine and L-homoalanine, were detected. As the second half of the cascade, the transamination step catalysed by ScALT, was examined before showing the ability of ScALT to produce the above amino acids (section 5.3.5), hence, the failure here must rely on the first step, the carboxylation of propanal catalysed by ScPDC. To demonstrate that the carbonate buffer is not a sufficient source of CO$_2$ for the carboxylation reaction and consequently the overall reaction, the enzymatic cascade was carried out in a pressurised CO$_2$ reactor (70 psi, 4.8 bar), in the same buffer (NaHCO$_3$-Na$_2$CO$_3$, 250 mM, pH 8.5) using acetaldehyde or propanal as the substrate (25 mM). After 48 hours, L-homoalanine (0.26 mM, from propanal) and L-alanine (0.04 mM, from acetaldehyde) were detected by HPLC using the FMOC-Cl derivatisation method (section 2.14.1 and Appendix A3). The carboxylation step does not to appear to procced in the carbonate buffer under atmospheric pressure, whilst the addition of CO$_2$ which saturates the solution, provides a supply of CO$_2$ (70 psi or 0.48 MPa, according to Henry's law $[CO_2^{aq}] = 162$ mM using equations in 2.12)

This can be explained by the differences in reactivity of different forms of CO$_2$. CO$_2$ is a stronger electrophile compared to its hydrated form (bicarbonate) which makes it a better substrate for the carboxylation reaction.[122]

In order to show that the carboxylation step and consequently the overall reaction were dependent on the CO$_2$ pressure, a set of reactions with varying CO$_2$ pressure were performed with propanal as the other substrate (1 mL final reaction volume). All other reaction conditions remained the same. Significant pressure dependencies of the L-homoalanine yield were observed (Figure 6-1). A 15% increase in the amount of L-homoalanine synthesised was detected by doubling the pressure (to 140 psi or 0.96 MPa).

Biocatalytic synthesis of the amino acid L-homoalanine by exploiting the CO2 fixation of an α-keto decarboxylase
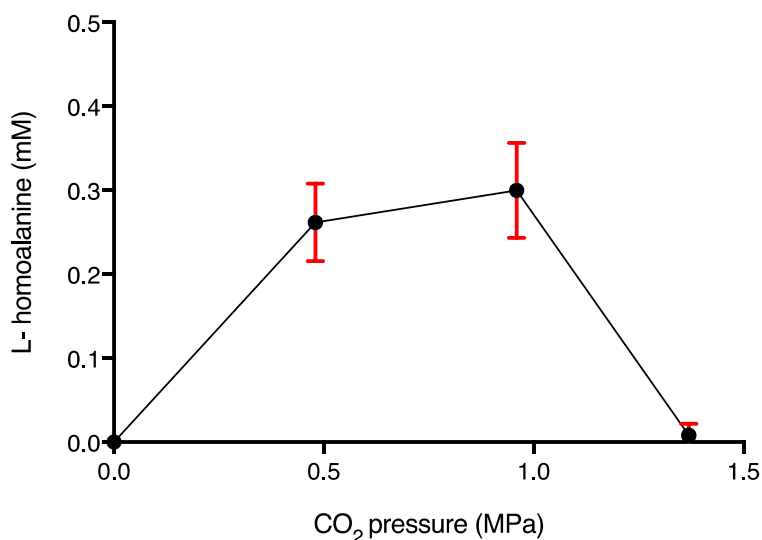
Figure 6-1. Effect of CO_2 pressure on the yield of L-homoalanine synthesis via the proposed biocatalytic approach. Error bars in red indicate standard deviations (SD) for three repeat experiments.

However, the improvement in the yield of L-homoalanine synthesis was found to be limited to within a certain pressure range. In fact, only a negligible amount of L-homoalanine was detected above pressure of 200 psi (or 1.37 MPa). This is possibly due to the inactivation of at least one of the enzymes resulting from dissolution of CO_2 in aqueous media forming carbonic acid and a drop in pH.[123] Additionally, transient carboxylation of nucleophilic residues such as Lys or His in the protein or possibly the C2 carbon of the ThDP could be possible affecting the overall yield. Deactivation of enzymes at high pressure was also observed by others in the enzymatic synthesis of L-methionine from methional and CO_2 using the reverse reaction of branched-chain decarboxylase (*Ll*kdcA) coupled with methionine aminotransferase (YbdL).[29] Pressure induced inactivation of *Sc*PDC was also reported elsewhere, showing that the enzyme lost its decarboxylase activity upon treatment of CO_2 at pressure of 145 psi (or 1 MPa) and above in either phosphate or MES buffer (100 mM, pH 6.0) at 35°C.[76] Accordingly, under pressure large oligomeric proteins dissociate into their subunit.[124] This is crucially important in the case of larger oligomeric proteins such *Sc*PDC where the enzyme must be at least a dimer to have catalytic activity.

6.2.2. Effect of pH on the yield of L-homoalanine synthesis

Next, the effect of pH on the reaction was evaluated. Each reaction was performed separately in three different buffers having different pH ranges: 250 mM sodium phosphate (NaH$_2$PO$_4$-

Biocatalytic synthesis of the amino acid L-homoalanine by exploiting the CO2 fixation of an α-keto decarboxylase

Na$_2$HPO$_4$, pH 7.0-8.0), 250 mM HEPES (pH 7.5) and 250 mM sodium carbonate (NaHCO$_3$-Na$_2$CO$_3$, pH 8.5-10.5). Each experiment was conducted inside a pressure reactor (1 ml) and the reaction was initiated by addition of CO$_2$ at the pressure of 70 psi (0.48 MPa). After 48 hours, the final pH of each reaction was determined with a narrow-range pH indicator strip and compared with the initial pH, it showed a drop of *ca.* one pH unit. This can be explained by the dissolution of the CO$_2$ into the reaction medium forming carbonic acid, an observation which has also been reported elsewhere.[29, 123]
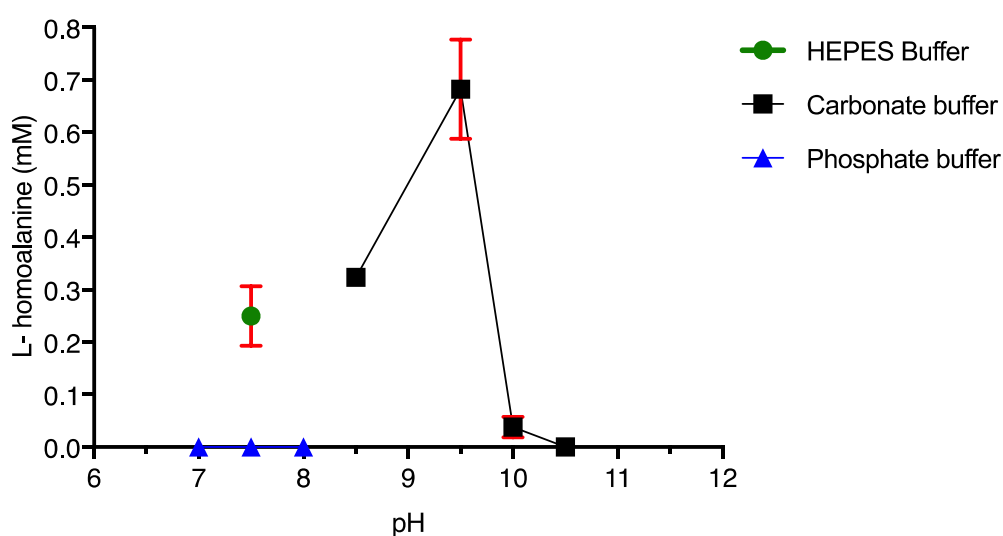
Figure 6-2. Effect of pH on the yield of L-homoalanine synthesis via the proposed biocatalytic approach. Error bars in red indicate standard deviations (SD) for three repeat experiments.

As can be seen from Figure 6-2, no product formation was observed when sodium phosphate buffer was used. L-homoalanine was synthesised when the buffer system was HEPES or carbonate. When HEPES was used to buffer the reaction, the concentration of L-homoalanine increased at pH 7.5. With carbonate buffer, a significant increase in concentration of L-homoalanine was detected. The improvement in the yield of L-homoalanine reached a maximum at an initial pH of 9.5 and then decreased significantly as the pH was increased to 10, with no product formation observed at pH 10.5.

However, the improvement in the carboxylation reaction by increasing the pH of the reaction cannot be viewed as a single parameter, because the pH changes also influenced the activity of the second enzyme, *Sc*ALT.

Since the coenzyme, ThDP plays a key role in the carboxylation step, it was necessary to test whether it can catalyse the reaction independently, in the absence of *Sc*PDC. Therefore, the reaction was performed as before using ThDP (1mM) on its own as the catalyst. The analysis of this reaction mixture after 48 hours did not show any detectable L-homoalanine (Figure 6-3). This indicates that although the coenzyme drives the chemistry of the reaction, it is unable to catalyse it in absence of the apoenzyme (*Sc*PDC).
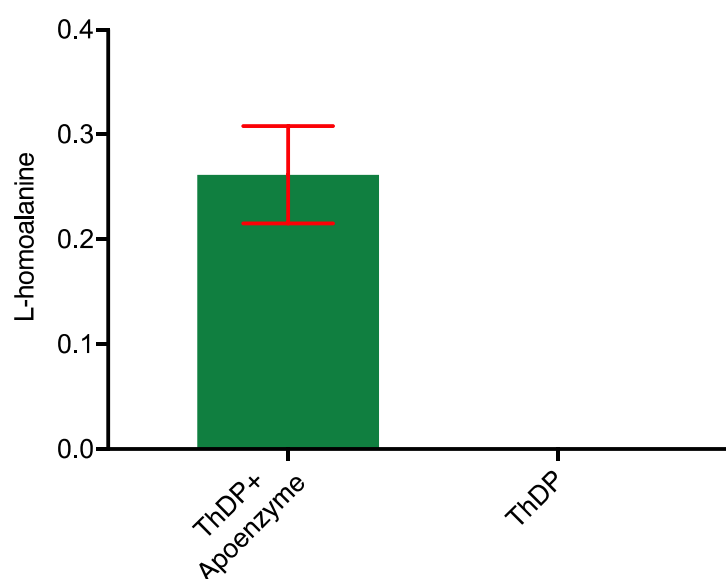
Figure 6-3. Investigation of the ability of the cofactor alone to catalyse the first step of the cascade. Error bars in red indicate standard deviations (SD) for three repeat experiments.

6.2.3. Time course of carboxylation-transamination

The carboxylation reaction of *Sc*PDC is a relatively slow process, being the reverse of the reaction the enzyme has evolved to achieve. For this reason, the first step in the proposed biocatalytic approach is limiting the overall efficiency to produce L-homoalanine. The kinetics of the reaction at the pressure of 70 psi (0.48 MPa) were monitored and analysed by taking samples at different intervals (0, 24, 48, 72, 96 and 120 hours) to find whether leaving the reaction for longer times would potentially improve the overall yield of the reaction or cause enzyme deactivation.
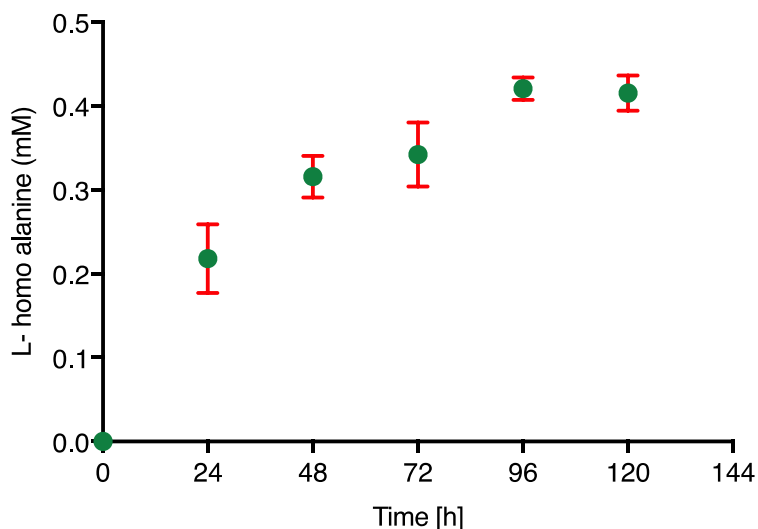
Figure 6-4. Time course of enzymatic synthesis of L-homoalanine following the proposed biocatalytic approach. Error bars in red indicate standard deviations (SD) for three repeat experiments.

A gradual increase in the concentration of L-homoalanine synthesised was observed as the reaction time increased with the maximum yield after 120 hrs (Figure 6-4). This observation was in good agreement with an earlier report,[47] where the carboxylation reaction of ScPDC was observed to be a rate limiting step in a multi-enzymatic system also involving lactate dehydrogenase for the production of L-lactic acid from ethanol and CO_2 via the formation of pyruvate as an intermediate. Elsewhere the reverse reaction of ScPDC, formation of pyruvate from acetaldehyde and CO_2, has been reported to be extremely slow.[80]

6.2.1. Effect of decarboxylase enzyme on the reaction

According to pH profile of LlKdcA, it was shown that LlKdcA has higher activity at alkaline pHs compared to ScPDC (Figure 3-15 and Figure 4-8). Additionally, the reverse reaction of LlKdcA was shown in combination with a transaminase enzyme (YbdL) for the production of L-methionine from methional and CO_2.[29] Hence, it was decided to investigate whether the replacement of the ScPDC with LlKdcA in our proposed enzymatic cascade (Scheme 6-3) could potentially improve the yield of the reaction or not. The reaction was performed in the carbonate buffer (250mM, pH 8.5) at the pressure of 70 psi (0.48 MPa) for 48 hours. All other reaction conditions remained the same. The only difference was the addition of LlKdcA instead of ScPDC to catalyse the carboxylation step.
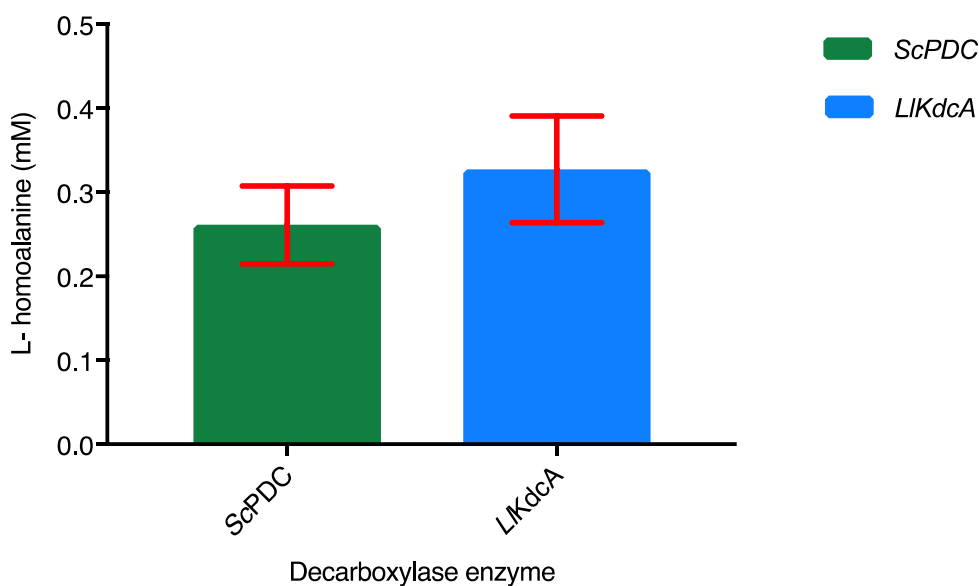
Figure 6-5. Effect of decarboxylase enzyme on the biocatalytic synthesis of L-homoalanine from propanal and CO_2. Error bars in red indicate standard deviations (SD) for three repeat experiments.

As can be seen from Figure 6-5, there was no significant difference in the yield of the L-homoalanine when *Ll*KdcA was applied as the carboxylase enzyme in the cascade. Therefore, the poor yield of the reaction could be due to the reasons other than the activity of the decarboxylase enzyme.

6.2.2. Investigation of the stability of the aldehyde, the substrate of the carboxylation

Despite of the considerable effort that was made to optimise the reaction by changing different variables as explained before, the reaction was still suffered from a poor yield (2.7% at pH 9.5). A low yield has also been reported in a similar carboxylation reaction using a decarboxylase for the synthesis of L- methionine from methional and CO_2 (29 psi,0.2 MPa) using the reverse reaction of *Ll*KdcA coupled with leucine dehydrogenase (LeuDH)(8% after 20 hours and 15% after 48 hours).[29] One reason for this is that aldehyde molecules, especially the short chain aliphatic ones, are very reactive and can undergo aldol-condensation reactions under either acidic or basic conditions in the presence of an inorganic salt such as carbonate salt (Scheme 6-4).[125, 126]

Therefore, the poor yield of the reaction could possibly be as a result of substrate degradation or conversion to the alternative aldol products. As this had not been investigated in the previous (*Sc*PDC/LDH)[46] and (*Ll*KdcA/LeuDH)[29] systems, it was decided to test this hypothesis by examining

the stability of the propanal in carbonate buffer using ^{1}H-NMR to monitor any composition changes.
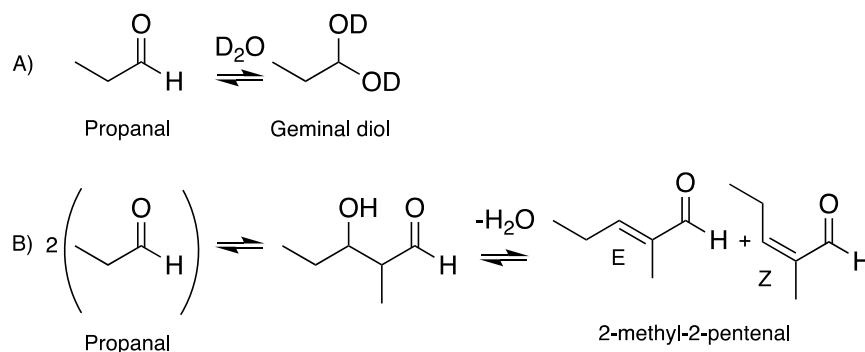
Scheme 6-4. The reactions of propanal in the carbonate buffer A) hydration B) Aldol-condensation

An aqueous solution of propanal (25 mM) in carbonate buffer (250 mM, pH 8.5) was prepared using deuterated water (The pD of 8.9 was calculated according to pD=pH+0.4 relation).[127] The stability of the propanal in the carbonate buffer was monitored by recording the ^{1}H NMR of the sample for 3 days at 24 hour intervals (0, 24, 48 and 72 hours). An equilibrium between the hydrated acetal (geminal-diol) (55%) and aldehyde (45%) forms of propanal was observed from the first ^{1}H NMR spectrum recorded in less than 5 min.[128] However, the percentages between these two species changed as the reaction proceeded. After 24 hours, the peaks corresponding to the product of self-condensation of propanal, 2-methyl-2-pentenal, catalysed by base (carbonate ion) were observed (Table 6-1).

Table 6-1. Percentages of different forms of propanal formed in the presence of carbonate ions

Time (minute)	Propanal (%)	Geminal-diol (%)	2-methyl-2-pentenal (%)
5	45	55	0
1440	43	43	14
2880	29	33	38
4320	18	26	56

The changes in the percentages of different forms of propanal were calculated by integrating the peak areas of each form showing that the self-condensation product was the major product after 72 hours. The ability of the carbonate ions to catalyse the aldol-condensation of propanal was

further confirmed by recording the ^{1}H NMR of the propanal in sodium phosphate (250mM, pH 8.5) buffer instead of carbonate. No aldol product was observed. The aldol-condensation could potentially be minimised by *in situ* generation of the aldehyde substrate, propanal. This could be achieved by the oxidation of the corresponding alcohol (propanal from propanol). This however requires an extra oxidation step, catalysed for example by an alcohol dehydrogenase enzyme and the addition of an expensive cofactor molecule, NAD[+]. Alternatively, the aldehyde substrate could be slowly released into the reaction environment from a more stable aldehyde adduct under the reaction conditions. The bisulfite adduct of aldehydes are a well-known form of aldehyde due to their enhanced stability.[129] They can be easily prepared from the reaction of alkali-metal bisulfite, usually sodium bisulfite, and aldehydes (Scheme 6-5).[129]
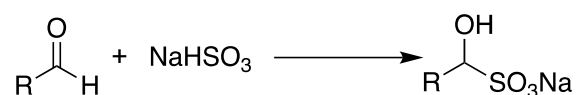
Scheme 6-5. Preparation of aldehyde-bisulfite adduct

The general procedure for recovery of aldehydes from their corresponding bisulfite adducts involves the treatment of the bisulfite adducts with a solution of sodium carbonate (1M).[129] More importantly, the decomposition of the aldehyde bisulfite back to its aldehyde is a pH dependent process with a better recovery at a higher pH of the carbonate buffer.[130, 131] Considering that the carboxylation reaction catalysed by a decarboxylase enzyme and consequently the overall reaction is favoured at the higher pH of carbonate buffer, the aldehyde-bisulfite could serve as a valuable precursor for the *in situ* generation of the aldehyde. Therefore, propanal bisulphite was synthesised by mixing of 1 equiv. of sodium bisulfite with 1 equiv. of propanal in methanol-water at room temperature.[132] After 3 days, a white solid was collected (92% yield). The analysis of the product by ^{1}H NMR showed peaks corresponding to the propanal bisulfite. The authenticity of the product was further confirmed by mass spectrometry showing a mass of 184.98 Da.

Next, propanal bisulfite (25 mM) was used instead of propanal in the proposed carboxylation-transamination cascade. The reaction was prepared in the carbonate buffer (250 mM, pH 9.5) and conducted inside a pressure reactor (1ml) at CO_2 pressure of 70 psi (0.48 MPa). After 48 hours, the reaction was analysed by HPLC and showed only negligible concentration of L-homoalanine formed (21.5 μM). The low yield of the reaction here can be explained by the slow release of the aldehyde into the reaction mixture. Therefore, there is still room for improvement
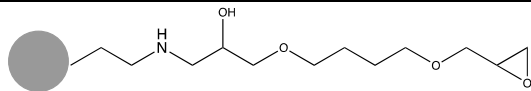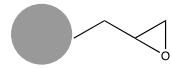
in the yield of the reaction, and this could likely to be achieved by increasing the time of the reaction. Additionally, increasing the concentration of the carbonate buffer could potentially increase the rate of regeneration of the aldehyde from the bisulfite adduct, which might eventually improve the overall yield of the reaction.

6.2.1. Immobilisation of the *Sc*PDC on the solid supports

As discussed earlier the carboxylation reaction catalysed by a decarboxylase enzyme, is a slow reaction. This means the reaction requires a longer time in order to proceed. Additionally, the enzyme needs pressurised CO_2 as the substrate. Incubation of the enzyme under the pressurized conditions for a long time would affect the stability of the enzyme. There are different approaches which can be applied to improve the stability of the enzyme at such conditions including the immobilization of the enzyme on a solid support.[76]

Here, the immobilisation of *Sc*PDC was carried on four different hydrophilic solid supports having different pore diameter and spacer size (Table 6-2). Relizyme supports are known to have larger pore size compared to Sepabeads which should favour better protein distribution and minimise protein 'caging' into narrow pores.[133]

Table 6-2. The typical physical and chemical features of the supports applied in this work.

Product Name	Functional group	Pore diameter (nm)
Sepabeads EC-HFA/S		10-20
Relizyme™ HFA403/S	Amino-epoxy	40-60
Relizyme™ EP113/S	epoxy	20-50
Relizyme™ EP403/S		40-60

The immobilisation was followed according to a procedure described previously, using cobalt (II) ions as the coordinating metal.[134] The immobilisation was performed by incubating the enzyme

with each resin between 4 and 24 hours. The binding of the enzyme on each support was assessed by measuring the decarboxylase activity after each incubation time (4 and 24 hours).
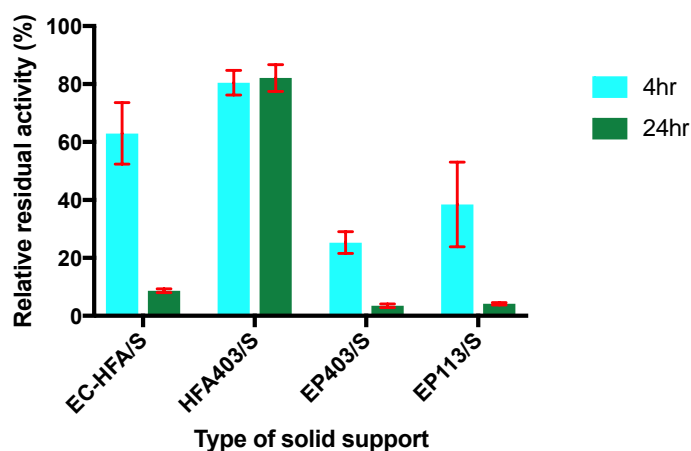
Figure 6-6.Immobilisation of *Sc*PDC on different resins monitored by measuring the decarboxylase activity of the enzyme and comparing it with the activity of the free enzyme. Error bars in red indicate standard deviations (SD) for three repeat experiments.

As can be seen from Figure 6-6, the HFA403/S resin offered a better support to anchor the enzyme as the activity of the immobilised enzyme was only 20% less compared to the activity of the free enzyme. Also, immobilisation was achieved after 4 hours with no significant change in the binding of the enzyme to the support after 24 hours. A fast immobilisation procedure as well as retained decarboxylase activity were advantage of HFA403/S resin over the others. HFA403/S resin with larger pore size in combination with a longer linker could possibly provide an even distribution of the enzyme on the surface of the support and prevent any aggregation. The *Sc*PDC-HFA403/S immobilised enzyme was stored at -20 ℃ prior to testing its application in the enzymatic synthesis of L-homoalanine.

6.3. Conclusion

In summary, a one-pot biocatalytic approach for the synthesis of L-homoalanine has been developed as an alternative route to the chemical synthesis. This is the first example of a non-proteinogenic amino acid being synthesized from an aldehyde and CO_2, implementing the reverse reaction of a decarboxylase enzyme, *Sc*PDC, coupled with a transamination enzyme, *Sc*ALT, which essentially assisted to displace the equilibrium position of the first step carboxylation. Our findings show that the carboxylation reaction and consequently the coupled reaction can only proceed when pressurized CO_2 is supplied. This highlights that carbonate buffer

alone is not enough to achieve the reverse reaction of ScPDC. The improvement in the yield of the reaction by the increase in the pressure of the CO_2 is limited to a certain range. The reaction was also attempted with another decarboxylase enzyme, LlKdcA, with no substantial improvement in the yield of the reaction. The pH dependency of the reaction has been established showing a maximum yield of 2.7% at pH 9.5. This pH is required for the release of the 2-ketoacid intermediate formed in the first step from the ScPDC cofactor (ThDP). The poor yield of the reaction (2.7%) has been attributed to the aldol-condensation of the aldehyde substrate, propanal, as monitored by ^{1}H NMR. This side reaction depletes the starting concentration of the aldehyde substrate, propanal, available for the biocatalytic cascade.

Our findings show that the substrate stability and its availability for the enzyme is one of the key factors that influence the reaction and its yield. In this context, engineering of the pressure reactor and equipping it with injection valves to gradually add the substrate to the system, could potentially prevents substrate degradation. However, this is beyond the scope and timeline of this project and it requires system design and manufacturing which on its own is a time-consuming process. Also, *in situ* generation of propanal is an alternative to minimise the self-condensation reaction and improve the yield of the L-homoalanine synthesis. Propanal bisulfite was synthesised and used in the coupled reaction in order to generate propanal *in situ* and minimise the aldol-condensation. However, the reaction still suffers from a poor yield. The future efforts in order to improve the yield of the reaction should be targeted at increasing the reaction time. However, the long reaction time may hamper the practical application of the proposed cascade on the industrial scale. Finally, enzyme engineering could be employed to improve the reaction yield by altering the specificity and efficiency of the enzyme. The expansion of the substrate range of the enzyme by protein engineering could make our proposed cascade a versatile route towards more proteinogenic and non-proteinogenic amino acid synthesis from CO_2.

Chapter 7. Rational protein design on *Sc*PDC

The application of enzymes as natural catalysts in chemical synthesis is often limited due to their narrow substate scope.[135] Herein, the expansion of the substrate range of the decarboxylase reaction of *Sc*PDC is of interest to increase the spectrum of aldehydes for the reverse activity of the enzyme. If the respective 2-keto acid is a substrate for decarboxylation, the binding of the corresponding aldehyde to the C2-carbon of the ThDP located at the active site for carboxylation should also be possible. This means that the aldehyde may be a possible substrate for the CO_2 fixation reaction.

Since high resolution X-ray crystal of structures of the *Sc*PDC are available,[44, 67] rational protein redesign by site-directed mutagenesis is a useful approach to create active site mutants of *Sc*PDC which could possibly accommodate larger substrates in their active site for both carboxylation and decarboxylation reactions. In this chapter, a summary of the key residues within the active site of *Sc*PDC is introduced. Furthermore, the active site mutants of both *Sc*PDC and *Zm*PDC (PDC from *Zymomonas mobilis*) which have been reported so far are introduced and discussed. In the light of previous studies and available structural data, new active site variants of *Sc*PDC which potentially have improved enzymatic properties with larger 2-keto acids have been designed, expressed, and evaluated. The kinetic parameters of these mutants have been measured and compared in two different buffers: sodium phosphate and citrate.

7.1. Background

Rational protein design by sited-directed mutagenesis (SDM) is a powerful technique which can be utilised to modify enzymes with improved enzymatic properties. As its name suggests, SDM allows for the specific substitution, insertion or deletion of any coded amino acid.[136] Oligonucleotide-directed mutagenesis is a commonly used technique *in vitro* to deliver the desired mutation at a specific position. The desired mutation is directed at the target site through a PCR reaction by amplifying the template using a pair of mutagenic primers. Following phosphorylation and ligation of the two ends of the newly synthesised DNA by T4-polynucleotide kinase and T4-ligase, a transformation reaction takes place in which the newly synthesised DNA is inserted into a host bacterial strain. DpnI digestion is performed prior to the transformation in order to digest the parental DNA and minimise the number of WT transformants (Figure 7-1).[50]

SDM has become a common tool for delineating reaction mechanisms. The method has also been applied to change substrate and cofactor specificity of enzymes.[137-139] However, protein redesign by SDM requires detailed knowledge of the structure of the enzyme and its reaction mechanism. The alternative is to generate a library of mutants and rely on high throughput activity screening, ideally linking genotype and phenotype in a directed evolution process.[55]
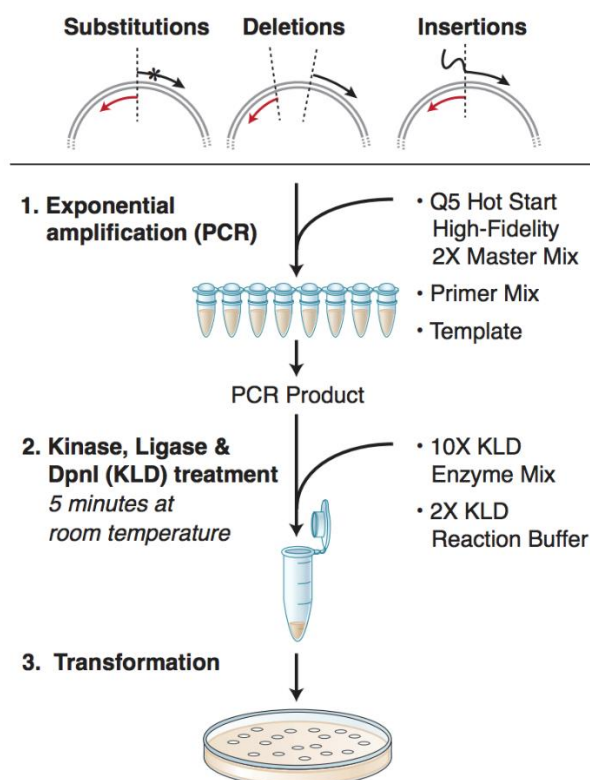
Figure 7-1. A graphical representation of the procedure used in the Q5® Site-Directed Mutagenesis kit (NEB). Step 1. Amplification of the DNA template by DNA polymerase using a mutagenic primer. Step 2. Treatment of the PCR product with kinase, ligase and DpnI (KLD) enzymes. Step 3. Transformation of the newly synthesised DNA product which carries the desired mutation.[140]

7.2. Design of the active site mutants of *Sc*PDC

7.2.1. Comparison of the *Sc*PDC and *Zm*PDC

Both *Sc*PDC and *Zm*PDC are homotetramers and require the cofactors thiamine diphosphate and Mg^{2+} for their catalytic activity.[141-144] Like *Sc*PDC, each subunit of *Zm*PDC consists of three domains (α, ß and γ). The active site, where the cofactors ThDP and Mg^{2+} are also present, is located between the first domain (α) of the one subunit and the third domain (γ) of the other subunit within a dimer ($\alpha_1\gamma_2$ and $\alpha_2\gamma_1$). The sequence alignment of *Sc*PDC (PDB accession number:

2VK1)[67] and *Zm*PDC (PDB accession number: 2WVA)[142] shows that the proteins share about 46 % sequence similarity (Figure 7-2).

```
ScPDC     MSEITLGKYLFERLKQVNVNTVFGLPGDFNLSLLDKIYEVEGMRWAGNANELNAAYAADG        60
ZmPDC     -MSYTVGTYLAERLVQIGLKHHFAVAGDYNLVLLDNLLLNKNMEQVYCCNELNCGFSAEG        59


ScPDC     YARIKGMSCIITTFGVGELSALNGIAGSYAEHVGVLHVVGVPSISAQAKQLLLHHTLGNG       120
ZmPDC     YARAKGAAAAVVTYSVGALSAFDAIGGAYAENLPVILISGAPNNNDHAAGHVLHHALGKT       119


ScPDC     DFTVFHRMSANISETTAMITDIATAPAEIDRCIRTTYVTQRPVYLGLPANLVDLNVPAKL       180
ZmPDC     DYHYQLEMAKNITAAAEAIYTPEEAPAKIDHVIKTALREKKPVYLEIACNIASMPCAAPG       179


ScPDC     LQTPIDMSLKPNDAESEKEVIDTILVLDKDAKNPVILADACCSRHDVKAETKKLIDLTQF       240
ZmPDC     PASALFNDEASDEASLNAAVEETLKFIANRDKVAVLVGSKLRA-AGAEEAAVKFADALGG       238


ScPDC     PAFVTPMGKGSIDEQHPRYGGVYVGTLSKPEVKEAVESADLILSVGALLSDFNTGSFSYS       300
ZmPDC     AVATMAAAKSFFPEENPHYIGTSWGEVSYPGVEKTMKEADAVIALAPVFNDYSTTGWTDI       298


ScPDC     YKTKNIVEFHSDHMKIRNATFPGVQMKFVLQKLLTAIADAAKG---YKPVAVPARTPANA       357
ZmPDC     PDPKKLVLAEPRSVVVNGIRFPSVHLKDYLTRLAQKVSKKTGALDFFKSLNAG-ELKKAA       357


ScPDC     AVPASTPLKQEWMWNQLGNFLQEGDVVIAETGTSAFGINQTTFPNNTYGISQVLWGSIGF       417
ZmPDC     PADPSAPLVNAEIARQVEALLTPNTTVIAETGDSWFNAQRMKLPNGARVEYEMQWGHIGW       417


ScPDC     TTGATLGAAFAAEEIDPKKRVILFIGDGSLQLTVQEISTMIRWGLKPYLFVLNNDGYTIE       477
ZmPDC     SVPAAFGYAV----GAPERRNILMVGDGSFQLTAQEVAQMVRLKLPVIIFLINNYGYTIE       473


ScPDC     KLIHGPKAQYNEIQGWDHLSLLPTFGAKD-YETHR-----VATTGEWDKLTQDKSFNDNS       531
ZmPDC     VMIHD--GPYNNIKNWDYAGLMEVFNGNGGYDSGAGKGLKAKTGGELAEAIK-VALANTD       530


ScPDC     KIRMIEVMLPVFDAPQNLVEQAKLTAATNAKQ------          563
ZmPDC     GPTLIECFIGREDCTEELVKWGKRVAAANSRKPVNKLL        568
```

Figure 7-2. Sequence alignment of PDC from *Sc*PDC and *Zm*PDC. The alignment was performed using FASTA sequence from PDB ID: 2VK1 (*Sc*PDC)[67] and 2WVA (*Zm*PDC)[142] and using the online sequence alignment tool (Clustal Omega). The ThDP binding motifs are highlighted in yellow, active site hydrophobic residues in turquoise and hydrophilic residues in green. The modified residues in this thesis are underlined.

Although the two enzymes only share 46% sequence similarity, the comparison of the crystal structure of both enzymes show that the active centre of PDC is lined with mainly hydrophobic and hydrophilic residues (Figure 7-3, A and B).[145]
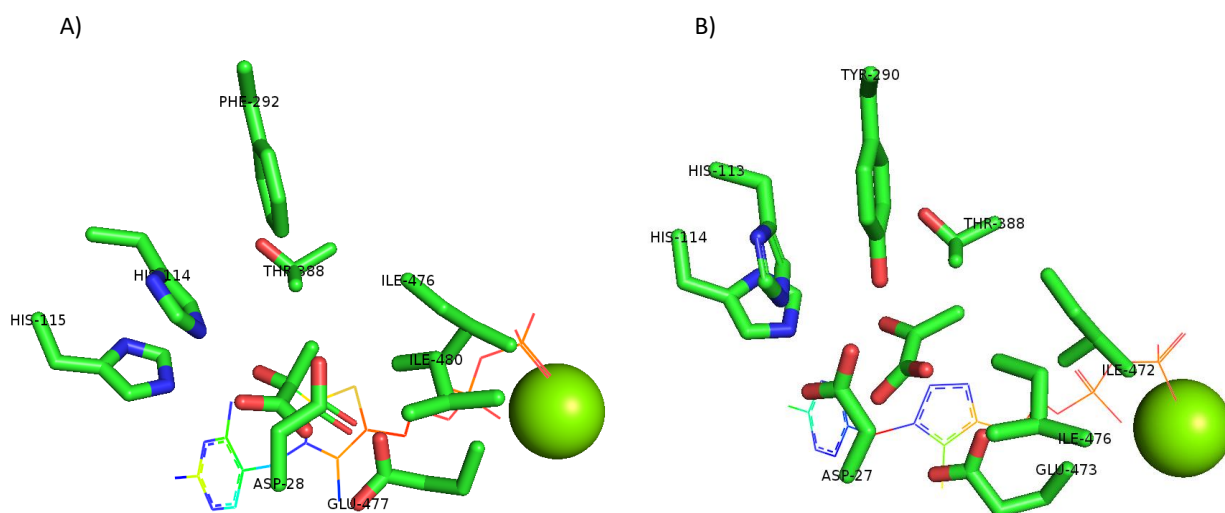
Figure 7-3. Comparison of the active site residues of the PDC from (A) *Saccharomyces cerevisiae* (PDBID:2VK1)[67] and (B) *Zymomonas mobilis* (PDBID:2WVA)[142]. The ThDP is shown as a wire frame structure and magnesium ion as a sphere. The figure was adapted from Andrews *et al.*[146] and recreated by PyMOL.

7.2.2. Hydrophilic residues in the active centre of PDC

There is a broad similarity in the residues which keep the substrate in the active site of *Sc*PDC and *Zm*PDC. Generally, there are four ionizable active site residues found in both enzymes (Table 7-1).[145]

Table 7-1. Hydrophilic amino acid residues involved in the wall of the cavity forming the active site in *Sc*PDC and *Zm*PDC. Residues studied by site-directed mutagenesis are marked with asterisks.

	*Sc*PDC	*Zm*PDC
Hydrophilic residues towards the pyrimidine-ring side	E477*	E473*
	D28*	D27*
	H114	H113*
	H115	H114*

The decarboxylation reaction of 2-ketoacids catalysed by PDC requires various protonation and deprotonation steps which underlines the importance of the presence of acid-base groups near the active site. Looking at the X-ray crystal structures of *Sc*PDC and *Zm*PDC shows that, in addition to the amino group of the 4'-aminopyridmidine ring of ThDP, the side chain of the above residues

(Table 7-1) are within a striking distance to the C2 atom of the thiazolium ring (Figure 7-4).[147] These residues are potentially involved in converting the thiazolium C2H to the C2 carbanion/ylide to initiate the decarboxylation reaction.[147]

A)

B)
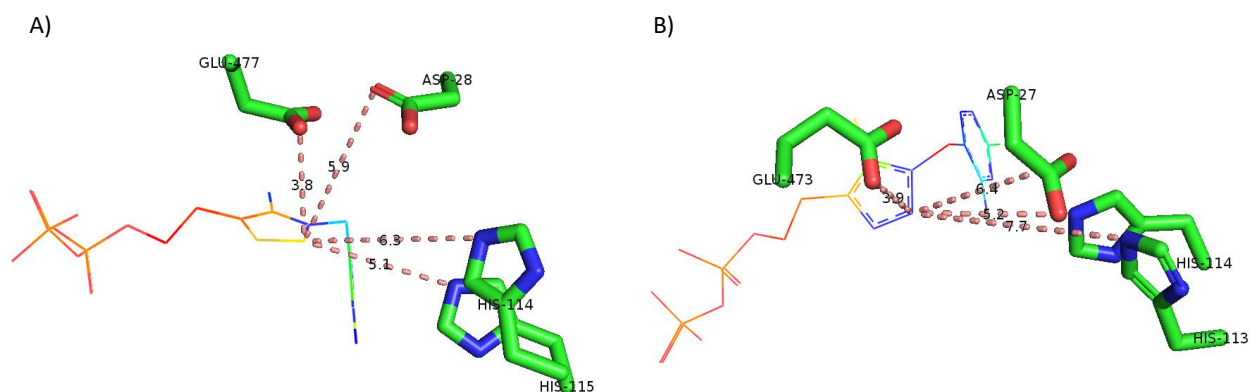
Figure 7-4. The distance between the hydrophilic amino acid residues and the C2 atom of ThDP. Measurements were made using crystal structures of A) ScPDC (PDBID:2VK1)[67] and B) ZmPDC (PDBID:2WVA)[142] by PyMoL. The ThDP is shown as a wire frame and distance in dashes.

Some of the ionizable active site residues in both ScPDC and ZmPDC have previously been subjected to mutagenesis studies.[145, 147-151] In both enzymes, mutations had a larger effect on k_{cat} compared to K_m values, suggesting that the ionizable residues are more important during catalysis rather than in substrate binding.[146] Glutamate residue at position 477 in ScPDC (E477) (corresponding to E504 in this work and E473 in ZmPDC) is one of the key hydrophilic amino acid residues which is in a close distance (3.8 A°) to the C2 carbon atom of the thiazolium ring of ThDP (Figure 7-5).
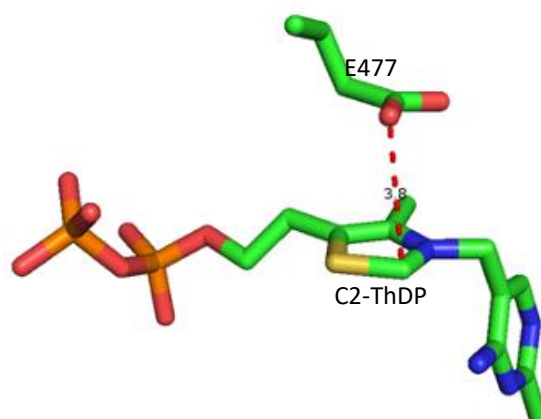
Figure 7-5. Location of residue E477 near ThDP in ScPDC. The figure was created by PyMOL using ScPDC (PDBID:2VK1)[67] crystal structure.

116

The importance of this glutamic acid residue has been previously studied in both *Sc*PDC and *Zm*PDC by site-directed mutagenesis. In *Sc*PDC, a substitution to aspartic acid (E477D) resulted in a decarboxylase activity of less than 1%.[152] Similarly low activity has been obtained when this residue was replaced with glutamine (E477Q).[147, 150] Therefore, both shortening the side chain of E477, and also the replacement of it with a non-ionizable group results in the loss of activity. These results underline the importance of this acidic residue in the catalysis.

Likewise, substitution of the equivalent a highly conserved glutamate residue at position 473 in *Zm*PDC with aspartic acid (E473D) and asparagine (E473N) resulted in the loss of decarboxylase activity (Table 7-2).[145]

Table 7-2. Comparison of decarboxylase activity of WT vs mutants of *Sc*PDC and *Zm*PDC

	*Sc*PDC	Decarboxylase activity [U/mg]	*Zm*PDC	Decarboxylase activity [U/mg]
Active site mutants of *Sc*PDC and its counterpart in *Zm*PDC targeted at the hydrophilic residues	WT	35-60	WT	120-150
	E477D	<1 %	E473D	0
	E477Q	<2 %	E473N	0

Herein, the E477Q (E504Q) mutant of *Sc*PDC was constructed and purified as this mutant can serve as a negative control in the decarboxylation reaction. It is of interest to us to examine the effect of this mutation later on in the reverse direction, i.e. carboxylation.

7.2.3. Hydrophobic residues in the active centre of PDC

Looking closer into the active site of *Sc*PDC, there are four hydrophobic residues which form the hydrophobic pocket surrounding the side chain of the substrate, pyruvate (Figure 7-3, A). These residues are known as 'substrate specificity' residues.[146] In *Zm*PDC, similar residues form the hydrophobic pocket around the side chain of the pyruvate (Figure 7-3, B). The only difference is, in *Zm*PDC, the phenylalanine (F292) is replaced by tyrosine (Y290) which forms an additional hydrogen bond with the carbonyl of the pyruvate.

Table 7-3. Comparison of hydrophobic residues which are involved in the cavity leading to the active site in *Sc*PDC and *Zm*PDC respectively. Residues previously studied SDM are marked with asterisks.

	*Sc*PDC	*Zm*PDC
Hydrophobic residues (Substrate specificity residues)	I476	I472*
	I480	I476*
	T388	T388
	F292	Y290

The I480 in yeast PDC (corresponding to I507 in this work and I476 in *Zm*PDC) is one of the conserved hydrophobic residues within the active site among the pyruvate decarboxylases. The side chain of this residue is located about 6.4 Å (0.64 nm) from the C2-ThDP in the active site and is considered to determine the substrate range of the decarboxylation reaction (Figure 7-6).[145]
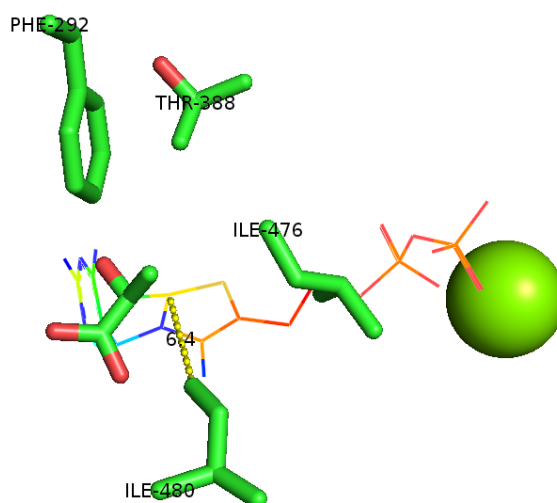
Figure 7-6. Hydrophobic residues in the active sites of *Sc*PDC. The distance of the I480 to the C2 carbon of ThDP is shown in dashes. The figure was created by PyMOL.

The influence of this isoleucine residue on the substrate range of the decarboxylation reaction has been previously studied in *Zm*PDC by SDM.[145] The study showed that the enlargement of the active site of *Zm*PDC by reducing the size of the side chain I476 and replacement of this residue with alanine (I476A), resulted in the loss of decarboxylase activity using pyruvate as the substrate. This was explained by the potential influence of I476 on the transition states of C2-

ThDP during the reaction.[145] Since the role of this residue (I480) in ScPDC has still remained unexplored, it was decided to modify this residue using the same approach taken for ZmPDC.

Looking at the crystal structure of the ScPDC (Figure 7-6), there are other hydrophobic residues where their modification to a smaller residue could potentially improve the substrate range of the enzyme by increasing the size of the active site. Thus, F292 (F319 in this work) and T388 (T415 in this work) residues were selected to be modified by SDM.

Although ScPDC and ZmPDC have similar residues that are in direct contact with the substrate pyruvate, ScPDC has much broader substrate range than ZmPDC.[153, 154] This is because in the active site of ScPDC, instead of a bulky tryptophan (W392), there is a small alanine residue (A392). Therefore, in this work, we have chosen ScPDC for modification by SDM in order to improve the decarboxylase activity of the enzyme with larger 2-keto acids. These mutants potentially could be applied in the carboxylation reaction of ScPDC for CO_2 fixation.

7.3. Results and discussion

7.3.1. Construction, overexpression, purification, and characterisation of ScPDC mutants

As the aim of this work was to expand the substrate scope of ScPDC in both decarboxylation and carboxylation reaction, those amino acids residues that were in proximity to the methyl group of the pyruvate in the crystal structure (Figure 7-3, A) were targeted for modification by SDM. These residues were E477, I480 and T388 in the original sequence correspond to E504, I507 and T415 in this work. The shift in the numbering was due to the presence of the two tags, His tag and AviTag ™,[69] followed by a thrombin cleavage site in the cloned ScPDC plasmid which was constructed in this work. The two tags were designed into the sequence in order to facilitate protein purification from E. coli. His Tagged proteins can be simply purified by the IMAC purification method. The lysine residue within the AviTag can be used to covalently attach biotin to the protein. The biotinylated proteins can be also purified using avidin or streptavidin.

New active site mutants of ScPDC (E504Q, I507A, T415S, T415A and F319A) were generated by SDM using a Q5® Site-Directed Mutagenesis kit (NEB) and a pair of forward/reverse primers (Table 2-12). For each mutation, the standard procedure of site directed mutagenesis was performed to substitute the targeted amino acid at the specified position with the desired one.

The generation of each mutant was driven by a PCR reaction which generally comprised of three steps as explained in section 2.9. The only difference between each mutation was the primers used (Table 2-12) and the annealing temperatures. These temperatures are summarized in Table 7-4.

Table 7-4. Annealing temperature applied in the PCR reaction for each mutation.

Mutation	PDC_E504Q	PDC_I507A	PDC_T415S	PDC_T415A	PDC_F319A
T_a (°C)	57	58	60	66	62

The success of each PCR reaction was examined by running the PCR product on 1% (w/v) agarose gel. This showed a bright band at 5.8 kb confirming that the amplification was successful (Figure 7-7).
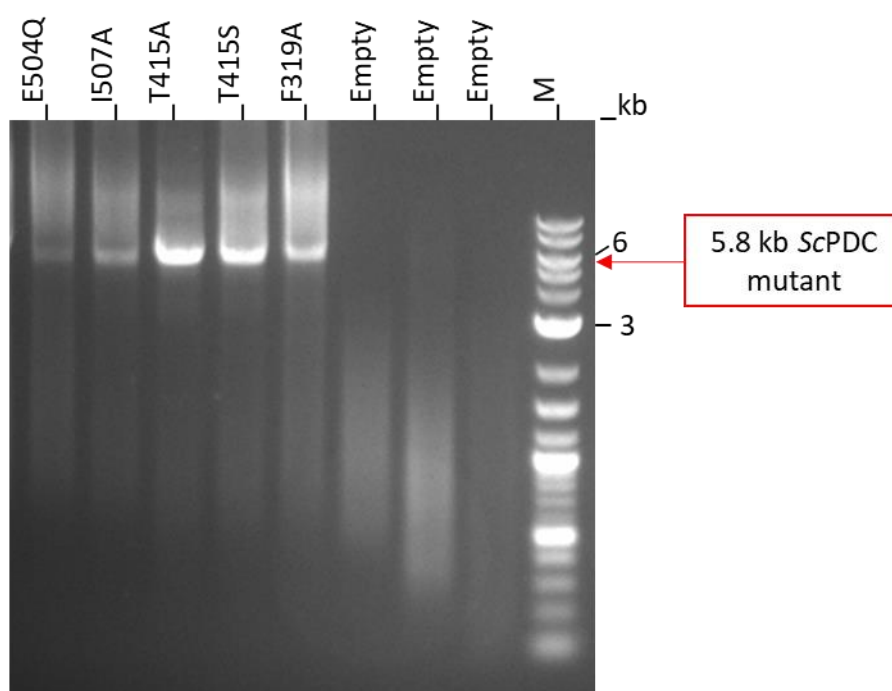
Figure 7-7. Agarose gel electrophoresis analysis of PCR fragments. Th PCR amplification was done using Q-5 hot start Master Mix and mutagenesis forward/reverse primers. M represent the DNA ladder (1kb Plus DNA Ladder, N0550G)

The PCR derived mutant plasmid was treated with KLD enzymes to circularize fragments and remove the original template DNA. After the KLD treatment, each mutant was transformed into *E. coli* DH5α chemically competent cells. The cell colonies were picked and grown overnight followed by plasmid extraction using MiniPrep extraction kit (Promega) to obtain the plasmid.

The correct sequence of each mutant was confirmed by DNA sequencing. Each mutant was then re-transformed into *E. coli* BL21(DE3) chemically competent cells and the cells were grown, and protein expressed using the same method as the wild type enzyme (section 3.2.2).

All mutants were purified using the nickel (II)-IMAC standard procedure. After sonication, the soluble fraction was loaded on a nickel affinity column, washed with buffer (50 mM sodium citrate, 500 mM NaCl, pH 6.5) and then the protein eluted with increasing concentration of imidazole from (0 to 1.0 M). The eluted fractions were examined by SDS-PAGE (Figure 7-8). The fractions containing *Sc*PDC were pulled together and dialyzed overnight against dialysis buffer (50 mM citrate, 2.0 mM $MgCl_2$, 1.0 mM ThDP, pH 6.5) to remove imidazole. The enzyme solution was concentrated and further purified using Vivaspin 20 centrifuge column (MWCO 30 kDa). Glycerol was added to the final concentration of 10% (v/v) and the protein was stored at $-80\,°C$.
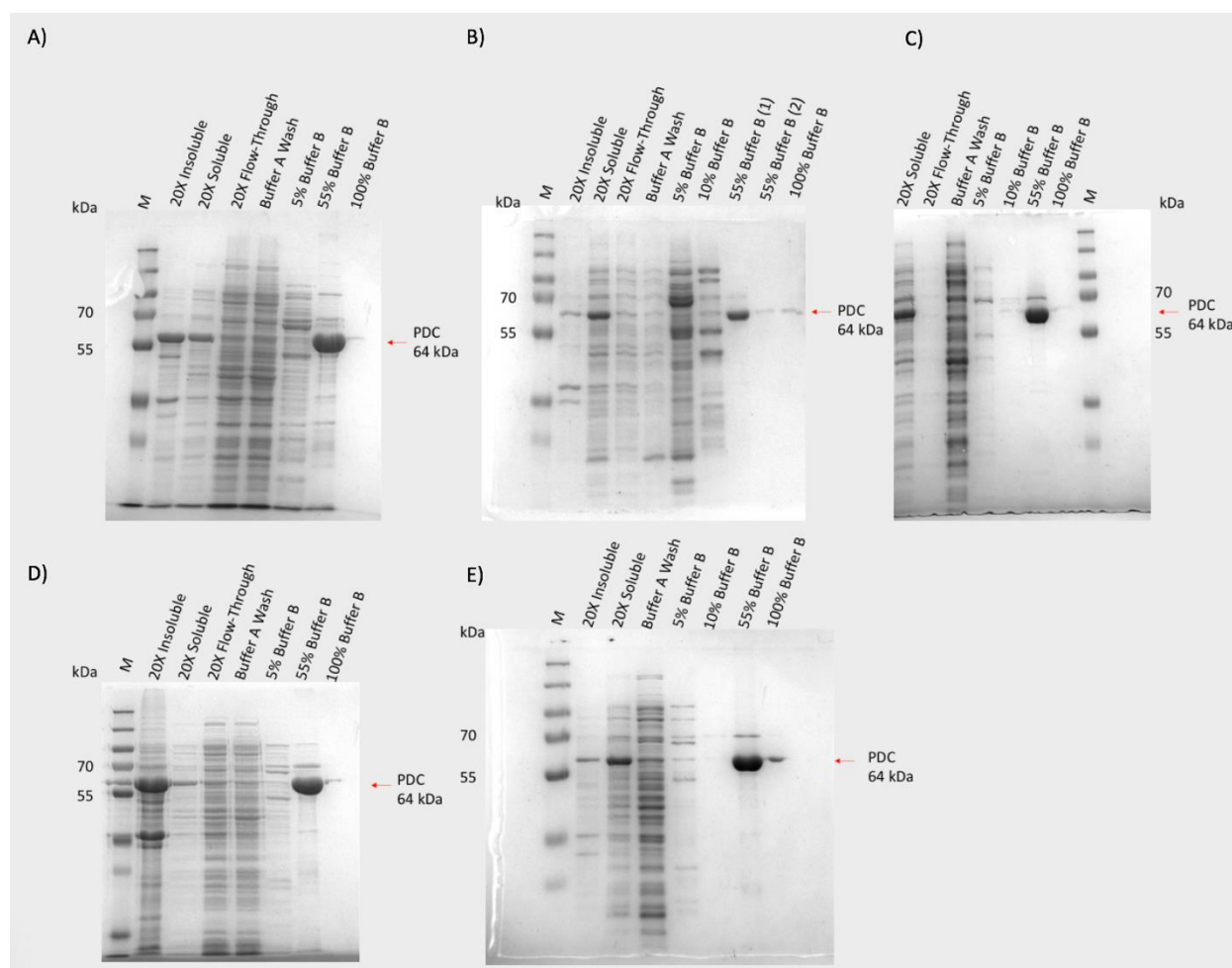
Figure 7-8. Purification of *Sc*PDC mutants analysed by SDS-PAGE gel. A) E404Q, B) I507A, C) F319A, D) T415S and E)T415A. M represents marker which is PageRuler[TM] Plus Prestained Protein Ladder

The His tag was retained on the mutants as well as the WT ScPDC as it has been reported that the presence and absence of the tag has no interference on the activity of the enzyme.[147]

Initially, the decarboxylase activity of the WT enzyme and the mutants were measured in sodium phosphate buffer (mM) at pH 6.0 at 30 ℃ with three aliphatic 2-keto acid substrates : Pyruvate **1** , 2-ketobutanoic acid **2** and 2-ketopentanoic acid **3** , each at the final concentration of 30 mM.[74]

As illustrated in Figure 7-9, the decarboxylase activity of the WT enzyme gets reduced as the length of the 2-keto acid substrate is increased from three to five carbons. The original activity of the WT enzyme was reduced by 36% with **2** and by 60% with **3**. The decreasing trend in the decarboxylase activity of ScPDC with substrate **2** and **3** was also published by Gocke.[74] Similar behaviour has also been reported for WT ZmPDC where the original decarboxylase activity of enzyme reduced by 34% with **2** and 90% with **3**.[155]
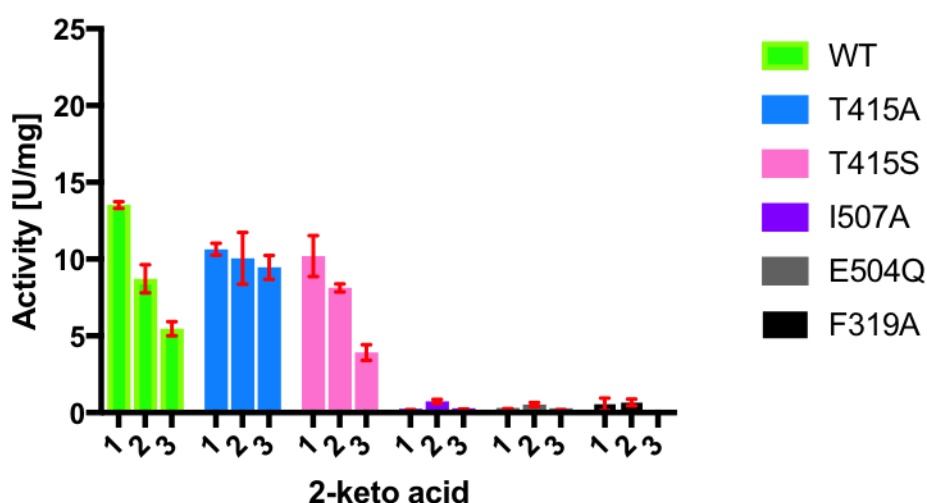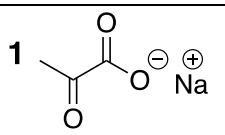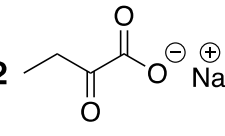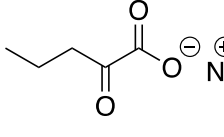
Figure 7-9. Specific decarboxylase activities of *Sc*PDC variants with aliphatic 2-keto acids. Results show 1: pyruvate, 2: 2-ketobutanoic acid and 3: 2-ketopentanoic acid. Measurements were made in sodium phosphate buffer pH 6.0 at 30 ℃ at a substrate concentration of 30 mM. Error bars in red indicate standard deviations (SD) for three repeat experiments.

The E504Q mutant has lost its decarboxylase activity with all three substrates: **1**, **2** and **3** as expected. This highlights the significance of this highly conserved glutamate residue in the decarboxylation reaction by ScPDC. The results here matched with the previous studies which have been reported for both ScPDC and ZmPDC.[145, 150]

The I507A and F319A mutants were designed and generated to expand the size of the active site to fit larger substrates. However, these mutations had a negative impact on the enzyme where

almost no decarboxylase activity was observed with all three substrates. It is important to compare the I507A mutant with its counterpart in *Zm*PDC (I476). In a study by Pohl *et al*, the role of the I476 in *Zm*PDC has been probed by generating three mutants of it: I476L, I476V and I476A with the aim of expansion in the size of the active site.[145] The results showed that the decarboxylase activity of the mutant enzymes was decreased in I476L and I476V to approximately 66% and 25% respectively using pyruvate as the substrate. The activity was almost non-existent in I476A. In summary, these results indicate the importance of the I507 and F319 residues in substrate recognition and binding, where its mutation to a small amino acid like alanine decreases the enzymatic activity. In the search for the active site mutants of *Sc*PDC with improved activity towards larger 2-ketoacids, the T415S mutant was generated. At first glance in Figure 7-9, T415S showed a similar trend of decrease in the decarboxylase activity with **1**, **2** and **3**. The decarboxylase activity of the T415S mutant decreased by 20% with **2** followed by 62% with **3** compared with pyruvate as the substrate (Table 7-5). These results were promising as the activity of the enzyme was not as bad as it was for the I507A and F319A mutants leading us to further investigate this residue. Therefore, it was decided to reduce the size of the side chain at position 415 further by changing it to alanine (T415A).

Table 7-5.Comparison of the substrate ranges of the decarboxylation reaction of WT, T415S and T415A. Measurements were made in sodium phosphate buffer at pH 6.00 at 30 °C at a substrate concentration of 30 mM. In parentheses are the percentage drop in the activity compared to the activity with the pyruvate.

2-keto acid	PDC specific activity [U/mg]		
	WT	T415S	T415A
1	13.5	10.2	10.6
2	8.72 (35)	8.13 (20)	10.1 (5)
3	5.47 (59)	3.92 (62)	9.76 (8)

Unlike the WT enzyme, where its decarboxylase activity is reduced as the size of the 2-ketoacid substrate is increased, the decarboxylase activity of T415A mutant almost remained the same

with all three substrates (Figure 7-9). The T415A mutant proved to have improved activity with both (**2**) and (**3**) compared to the WT enzyme showing 16% increased activity with (**2**) and 78% increased activity with (**3**). In order to understand the effect of each mutation on each variant, the kinetic parameters, including the substrate binding affinity (Michaelis constant) (K_m), catalytic efficiency (k_{cat}) and the ratio of k_{cat}/K_m often termed as the specificity constant, were determined for the WT, T415S and T415A with all three substrates in sodium phosphate buffer. The data was fitted to the Hill equations as described in section 2.8.

As observed in Figure 7-10, the WT enzyme exhibited a sigmoidal V/[S] plot as previously reported for ScPDC, confirming the allosteric activation of the enzyme by the substrate pyruvate.[143] Hill coefficient of greater than one, 3.17, indicates positive cooperativity.

PDC-WT

K_m = 7.61 mM
k_{cat} = 58.3 s^{-1}
n= 3.17
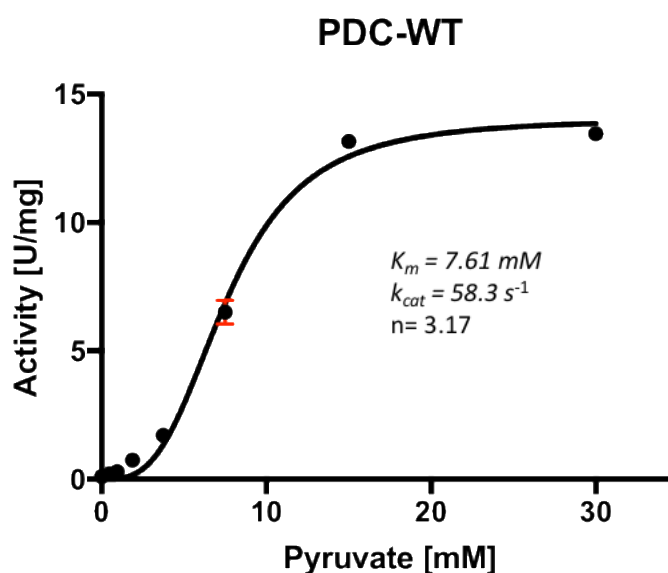
Figure 7-10. A plot of reaction rate versus pyruvate concentration for wild type ScPDC.

The kinetic parameters of the WT, T415S and T415A mutants are summarized in (Table 7-6) showing the alterations in the values of K_m, k_{cat} and the ratio of k_{cat}/K_m. An increase in the affinity of the T415A mutant for (**1**) was observed: K_m is 6.35 mM versus 7.61 mM for the WT enzyme. The increase in the affinity of the enzyme for the substrate is more substantial when (**2**) was used as the substrate: K_m is 2.78 mM versus 5.16 mM for the WT. Having a closer look at K_m values of (**2**), there is a decreasing trend: K_m is 5.16 mM for the WT enzyme, 4.25 mM for T415S and 2.78 mM for the T415A. This can be interpreted as the affinity of the enzyme for (**2**) is gradually improved with the gradual enlargement of the active site where the T415 is changed to serine and then to alanine.

Table 7-6.Comparison of the kinetic constants for the decarboxylase activity of *Sc*PDC variants. Reactions were carried out in the phosphate buffer at pH 6.0 at 30 °C. *n.d.*: not determined. In parentheses are the relative values expressed as a percentage of the WT.

Compound	Parameter	WT	T415S	T415A
Pyruvate (**1**)	K_m (mM)	7.61	*n.d.*	6.35
	k_{cat} (s^{-1})	58.3	44.0	45.9
	k_{cat}/K_m (mM^{-1} s^{-1})	7.38 (100)	*n.d.*	7.23 (98)
2-ketobutanoic acid (**2**)	K_m (mM)	5.16	4.25	2.78
	k_{cat} (s^{-1})	30.2	37.0	38.8
	k_{cat}/K_m (mM^{-1} s^{-1})	5.85 (100)	8.71 (149)	14.0 (239)
2-ketopentanoic acid (**3**)	K_m (mM)	3.92	7.74	4.44
	k_{cat} (s^{-1})	23.3	36.2	43.6
	k_{cat}/K_m (mM^{-1} s^{-1})	5.95 (100)	4.68 (79)	8.30 (139)

The above studies were repeated in the sodium citrate buffer as this buffer has been reported to be a standard assay buffer for the measurements of the PDC activity.[73] Generally, the WT enzyme as well as the two mutants (T415A and T415S) showed that they have higher decarboxylase activity in the citrate buffer compared to phosphate buffer (Figure 7-11).
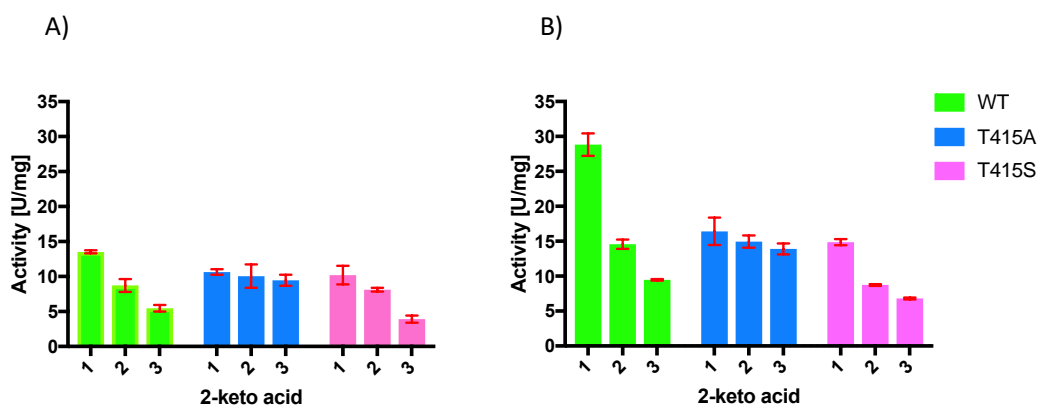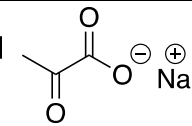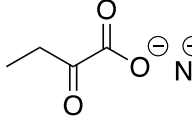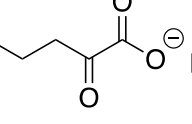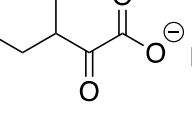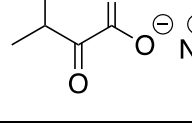
Figure 7-11. Comparison of the decarboxylase activity of WT, T415A and T415S using different substrates. Result show 1: pyruvate, 2: 2-ketobutanoic acid and 3: 2-ketopentanoic acid in two different buffers: A (sodium phosphate) and B (sodium citrate buffer). Error bars in red indicate standard deviations (SD) for three repeat experiments.

The decreasing trend in the activity of the WT and the T415S mutant with the increase in the size of the 2-keto acid was also apparent in the citrate buffer (Figure 7-11, B). The decarboxylase activity of the wild type enzyme reduced by 49% with (**2**) and 67% with (**3**) compared to activity with pyruvate as the substrate. Also, the T415S mutant showed a 41% drop in the decarboxylase activity with (**2**) and 54% with (**3**) (Table 7-7).

Table 7-7. Comparison of the substrate ranges of the decarboxylation reaction of WT, T415S and T415A. Measurement were made in sodium citrate buffer at pH 6.0 at 30 ℃ . In parentheses are the percentage drop in the activity compared to the activity with pyruvate.

2-keto acid	PDC specific activity [U/mg]		
	WT	T415S	T415A
1	28.8	14.9	16.4
2	14.6 (49)	8.72 (41)	14.9 (9)
3	9.46 (67)	6.80 (54)	13.9 (15)
4	0 (100)	0.130 (99)	0.770 (95)
5	9.36 (68)	6.09 (59)	5.66 (65)

The substrate range of the T415S and T415A mutants was furthered explored by testing their activities with branched chain substrate (**4**) and (**5**) using direct decarboxylase assay (section 2.7). For WT enzyme, similar results to those in literature were observed.[74] The enhancement in the activities of T415S and T415A mutants were limited to the straight chain substrates.

The kinetic parameters were also obtained in the citrate buffer for the WT as well as the two mutants (Table 7-8). As can be seen in Figure 7-12, in the absence of phosphate, the WT enzyme

has only a moderate degree of cooperativity: n_H of 1.32 in citrate versus 3.17 in phosphate buffer. This means that in the presence of the phosphate, the allostery of the enzyme is enhanced.

PDC-WT
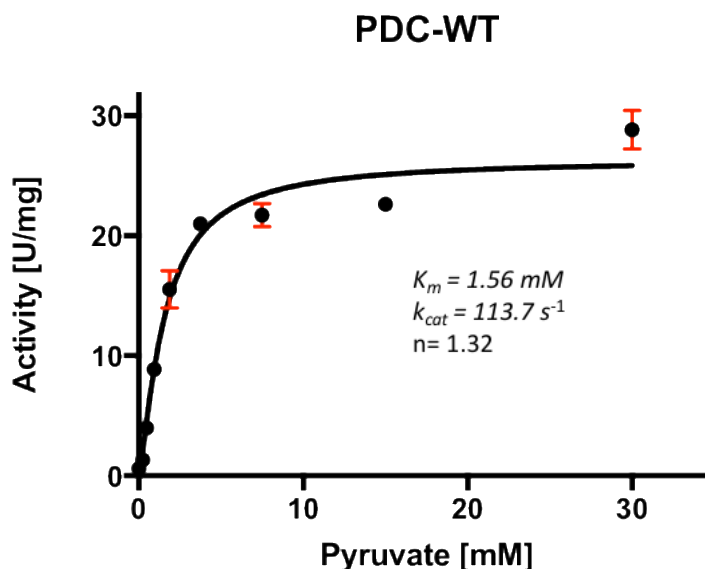
Figure 7-12. A plot of reaction rate versus pyruvate concentration for WT type *Sc*PDC in citrate buffer. Error bars in red indicate standard deviations (SD) for three repeat experiments.

The K_m value of 1.56 mM for the WT enzyme for pyruvate was obtained. This value was similar to data published by Holzer *et al*.[156] Although the enzyme was shown to have a lower degree of cooperativity in the citrate buffer, the affinity of enzyme is enhanced for the pyruvate in the absence of the phosphate. Furthermore, the K_m value was shifted from 7.61 mM in the phosphate buffer to 1.56 mM in citrate buffer for the WT enzyme. This can be explained by the inhibitory role of the phosphate on the enzyme, an observation which was also reported by Boiteux *et al*.[77] The inhibition of the enzyme by phosphate can be explained as the phosphate mimics the structure of the 2-keto acid carboxylate, therefore binding to the active site in the place of substrate. Looking at the K_m values, the increase in the affinity of the enzyme for the substrate in the citrate buffer was observed for the WT as well as the mutants with all three substrates (Table 7-8).

Table 7-8. Comparison of the kinetic constants for the decarboxylase activity of *Sc*PDC variants. Reactions were carried out in the citrate buffer at pH 6.00 at 30 °C as described in Material and methods. In parentheses are the relative values expressed as a percentage of the WT.

Compound	Parameter	WT	T415S	T415A
Pyruvate (**1**)	K_m (mM)	1.56	1.69	1.19
	k_{cat} (s^{-1})	113.7	65.7	69.6
	k_{cat}/K_m (mM^{-1} s^{-1})	72.9 (100)	38.9(53)	58.5 (80)
2-ketobutanoic acid (**2**)	K_m (mM)	0.655	0.598	0.597
	k_{cat} (s^{-1})	61.4	37.5	70.1
	k_{cat}/K_m (mM^{-1} s^{-1})	93 (100)	62.7 (67)	117.4 (126)
2-ketopentanoic acid (**3**)	K_m (mM)	0.739	1.25	0.924
	k_{cat} (s^{-1})	41.6	31.5	53.3
	k_{cat}/K_m (mM^{-1} s^{-1})	56.3 (100)	25.2 (45)	57.7 (102)

Looking at the kinetic data obtained in both phosphate (Table 7-6) and citrate buffer (Table 7-8), the T415A mutant showed to have improved specificity with both (**2**) and (**3**). This improvement was more significant when (**2**) was used as the substrate.

7.4. Conclusions

In summary, new active site mutants of *Sc*PDC were generated. Subsequent mutagenesis, aided by the crystal structure of *Sc*PDC in combination with the knowledge of the active site mutants of *Sc*PDC and *Zm*PDC that have been obtained so far, allowed for structure-guided modification of *Sc*PDC. From 5 different active site variants, T415A mutants has shown improved specificity towards larger 2-ketoacid substrates. The improvement in the specificity of this mutant with (**2**) has generated a useful biocatalyst not only in the decarboxylation direction but, potentially in the reverse reaction, i.e., carboxylation. The future work will aim to use this variant in our new biocatalytic approach for the synthesis of L-homoalanine, from propanal and carbon dioxide.

Chapter 8. Summary and Future work

In this work the reverse reaction of two decarboxylase enzymes, *Sc*PDC and *Ll*KdcA, have been studied in detail for the fixation of CO_2. Initially carbonate buffer has been used as a source of CO_2 to achieve the carboxylation reaction of each enzyme separately. Our results have showed that the carbonate buffer alone is not a sufficient source of CO_2 to achieve the reverse reaction of *Sc*PDC and *Ll*KdcA. Later, a one-pot enzymatic cascade has been designed involving two steps: carboxylation and transamination. The second step, transamination, has been incorporated not only to drive the equilibrium of the first step towards carboxylation but also to produce amino acids from CO_2 and aldehydes. The synthesis of L-homoalanine, from propanal and CO_2, has been investigated following our proposed cascade. The product of the reaction, L-homoalanine, is a key chemical platform for the synthesis of enantiopure active pharmaceutical agents. The cascade has been studied in depth in order to understand the parameters that directly influence the reaction. Testing the cascade in the carbonate buffer failed to generate any L-homoalanine. This confirmed that the carbonate buffer alone is not an enough CO_2 to drive the equilibrium of the first step towards carboxylation. In fact, the reaction can only succeed when pressurized CO_2 is supplied. Although pressurised CO_2 is needed to achieve the carboxylation and consequently the overall reaction, the reaction is limited to a certain pressure range. This limitation can be linked to the stability of each enzyme at such pressure. *Sc*PDC was also immobilised on different solid supports. From 4 different resins tested, the HFA403/S has provided a better support as the activity of the immobilised enzyme is similar to the free enzyme.

In addition to CO_2, the reaction is also pH dependent with a maximum yield at pH 9.5. This pH is required in order to facilitate the cleavage of the intermediate of the first step, carboxylation, from the cofactor, ThDP. The major limiting factor for the cascade going to competition has been identified as the competing aldol condensation of propanal that occurs under the preferred reaction conditions. One way to overcome this limitation is to generate propanal *in situ*. Accordingly, an adduct of propanal, propanal bisulfite, has been prepared as a more stable form of propanal. Also, the regeneration of propanal bisulfite back to propanal is a pH dependent reaction and occurs predominately at alkaline pHs. This condition is compatible with the reaction condition that favours the carboxylation reaction. Propanal bisulfite has been applied as the starting material instead of propanal in the cascade in the presence of carbonate buffer (250 mM, pH 9.5). However, the yield of L-homoalanine synthesised is lower compared to when

propanal was used. This could be due to the slow regeneration of propanal. According to our findings and the work has been done so far, the future work should focus on

- Start the cascade with propanal bisulfite instead of propanal and perform the reaction for longer time. This longer time could potentially improve the rate of regeneration of propanal from propanal bisulfite. Alternatively, higher concentration of carbonate buffer could be used to speed up the regeneration of propanal from propanal bisulfite.
- Application of engineered *Sc*PDC (T415A) instead of WT *Sc*PDC in the biocatalytic cascade. The kinetic studies of this mutant have showed that this mutant has improved catalytic efficiency towards 2-ketobutyric acid, the intermediate of the proposed cascade.
- Testing the immobilised *Sc*PDC on the HFA403/S support instead of free *Sc*PDC in the cascade. The immobilised enzyme could potentially be used in a flow system to continuously pump substrate into the reactor.
- Change the system from batch to continuous in order to feed the substrate, aldehyde, into the system and increase its basal concentration.
- Improve the equilibrium of the carboxylation reaction by removing the product from the reaction, for example by precipitation using quaternary ammonium salt.
- To use alternative coupling enzymes which have better equilibrium compared to transaminases, for example amino acid dehydrogenases (AADHs). AADHs are NADPH dependent enzymes that catalyse the amination of keto acids in the presence of ammonia. The equilibrium of AADHs is favourable towards the synthesis of amino acid.

References

1. J. Chapman, A. E. Ismail and C. Z. Dinu, *Catalysts*, 2018, **8**, 238.
2. K. Faber, in *Biotransformations in organic chemistry*, Springer- Verlag, Berlin, 5th edn., 2004, ch. 1. Introduction and background information, pp. 3-5.
3. K. M. Koeller and C. H. Wong, *Nature*, 2001, **409**, 232-240.
4. A. Radzicka and R. Wolfenden, *Science*, 1995, **267**, 90-93.
5. M. T. Reetz, *The Chemical Record*, 2016, **16**, 2449-2459.
6. J. M. Woodley, *Curr. Opin. Chem. Biol.*, 2013, **17**, 310-316.
7. M. Pawan, *Int. Res. J. Environ. Sci.*, 2014, **3**, 72-78.
8. E. J. Beckman, *Nature*, 2016, **531**, 180.
9. S. M. Glueck, S. Gumus, W. M. F. Fabian and K. Faber, *Chem. Soc. Rev.*, 2010, **39**, 313-328.
10. J. Micklefield, R. Lewin and M. Thompson, in *Science of Synthesis: Biocatalysis in Organic Synthesis II*, Georg Thieme Verlag, 2014, vol. 2, pp. 133-157.
11. M. D. Burkart, N. Hazari, C. L. Tway and E. L. Zeitler, *ACS Catal.*, 2019, **9**, 7937-7956.
12. C. B. Field, M. J. Behrenfeld, J. T. Randerson and P. Falkowski, *Science*, 1998, **281**, 237-240.
13. L. Schada von Borzyskowski, R. G. Rosenthal and T. J. Erb, *J. Biotechnol.*, 2013, **168**, 243-251.
14. H. G. Wood, S. W. Ragsdale and E. Pezacka, *Trends Biochem. Sci.*, 1986, **11**, 14-18.
15. M. Calvin, *Nature*, 1961, **192**, 799.
16. M. C. Evans, B. B. Buchanan and D. I. Arnon, *Proc. Natl. Acad. Sci. U. S. A*, 1966, **55**, 928-934.
17. L. G. Ljungdahl, *Annu. Rev. Microbiol.*, 1969, **23**, 515-538.
18. S. Herter, J. Farfsing, N. Gad'on, C. Rieder, W. Eisenreich, A. Bacher and G. Fuchs, *J. Bacteriol.*, 2001, **183**, 4305-4316.
19. I. A. Berg, D. Kockelkorn, W. Buckel and G. Fuchs, *Science*, 2007, **318**, 1782.
20. H. Huber, M. Gallenberger, U. Jahn, E. Eylert, I. A. Berg, D. Kockelkorn, W. Eisenreich and G. Fuchs, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 7851-7856.
21. T. J. Erb, I. A. Berg, V. Brecht, M. Müller, G. Fuchs and B. E. Alber, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 10631-10636.
22. G Schneider, a. Y Lindqvist and C. I. Branden, *Annu. Rev. Biophys. Biomol. Struct*, 1992, **21**, 119-143.
23. R. K. Thauer, *Science*, 2007, **318**, 1732.
24. A. Alissandratos and C. J. Easton, *Beilstein J. Org. Chem.*, 2015, **11**, 2370-2387.
25. S. Gutteridge and J. Pierce, *Proc. Natl. Acad. Sci.*, 2006, **103**, 7203-7204.
26. B. B. Buchanan and D. I. Arnon, *Photosynth. Res.*, 1990, **24**, 47-53.
27. I. A. Berg, D. Kockelkorn, W. Buckel and G. Fuchs, *Science*, 2007, **318**, 1782-1786.
28. U. Jahn, H. Huber, W. Eisenreich, M. Hugler and G. Fuchs, *J. Bacteriol.*, 2007, **189**, 4108-4119.
29. J. Martin, L. Eisoldt and A. Skerra, *Nat. Catal.*, 2018, **1**, 555-561.
30. J. R. Allen and S. A. Ensign, *J. Bacteriol.*, 1996, **178**, 1469-1472.
31. F. J. Small and S. A. Ensign, *J. Bacteriol.*, 1995, **177**, 6170-6175.
32. A. M. Krishnakumar, D. Sliwa, J. A. Endrizzi, E. S. Boyd, S. A. Ensign and J. W. Peters, *Microbiol. Mol. Biol. Rev.*, 2008, **72**, 445-456.
33. K. Kirimura, H. Gunji, R. Wakayama, T. Hattori and Y. Ishii, *Biochem. Biophys. Res. Commun.*, 2010, **394**, 279-284.
34. Y. Ishii, Y. Narimatsu, Y. Iwasaki, N. Arai, K. Kino and K. Kirimura, *Biochem. Biophys. Res. Commun.*, 2004, **324**, 611-620.
35. R. Santha, N. Appaji Rao and C. S. Vaidyanathan, *Biochem. Biophys. Acta*, 1996, **1293**, 191-200.
36. J. Ren, P. Yao, S. Yu, W. Dong, Q. Chen, J. Feng, Q. Wu and D. Zhu, *ACS Catal.*, 2016, **6**, 564-567.
37. M. Wieser, T. Yoshida and T. Nagasawa, *Tetrahedron Letters*, 1998, **39**, 4309-4310.
38. T. Yoshida and T. Nagasawa, *J. Biosci. Bioeng.*, 2000, **89**, 111-118.
39. M. Wieser, T. Yoshida and T. Nagasawa, *J. Mol. Catal. B: Enzym.*, 2001, **11**, 179-184.
40. M. Wieser, T. Yoshida and T. Nagasawa, *Tetrahedron Lett.*, 1998, **39**, 4309-4310.

41. H. Omura, M. Wieser and T. Nagasawa, *Eur. J. Biochem.*, 1998, **253**, 480-484.
42. T. Yoshida, K. Fujita and T. Nagasawa, *Biosci., Biotechnol., Biochem.*, 2002, **66**, 2388-2394.
43. K. A. Payne, S. A. Marshall, K. Fisher, M. J. Cliff, D. M. Cannas, C. Yan, D. J. Heyes, D. A. Parker, I. Larrosa and D. Leys, *ACS Catal.*, 2019, **9**, 2854-2865.
44. P. Arjunan, T. Umland, F. Dyda, S. Swaminathan, W. Furey, M. Sax, B. Farrenkopf, Y. Gao, D. Zhang and F. Jordan, *J. Mol. Biol.*, 1996, **256**, 590-600.
45. M. Miyazaki, M. Shibue, K. Ogino, H. Nakamura and H. Maeda, *Chem. Commun.*, 2001, 1800-1801.
46. M. Miyazaki, K. Ogino, M. Shibue, H. Nakamura and H. Maeda, *Chem. Lett.*, 2002, **31**, 758-759.
47. X. Tong, B. El-Zahab, X. Zhao, Y. Liu and P. Wang, *Biotechnol. Bioeng.*, 2011, **108**, 465-469.
48. D. Burke, D. Dawson and T. Stearns, Cold Spring Harbor Laboratory Press, 2000, pp. 113-114.
49. E. Haid, P. Lehmann and J. Ziegenhorn, *Clin. Chem.*, 1975, **21**, 884-887.
50. J. Dordick, in *Biocatalysts for Industry*, Plenum Press, New York, 1991, ch. 13.Protein Engineering of subtilisin, pp. 257-283.
51. S. M. G. Duff, T. J. Rydel, A. L. McClerren, W. L. Zhang, J. Y. Li, E. J. Sturman, C. Halls, S. Y. Chen, J. M. Zeng, J. X. Peng, C. N. Kretzler and A. Evdokimov, *Arch. Biochem. Biophys.*, 2012, **528**, 90-101.
52. E. K. Hanko, A. C. Paiva, M. Jonczyk, M. Abbott, N. P. Minton and N. Malys, *Nat. Commun.*, 2020, **11**, 1-14.
53. C. Neuberg and L. Karczag, *Biochem. Z.*, 1911, **36**, 68-75.
54. J. M. Candy, R. G. Duggleby and J. S. Mattick, *J. Gen. Microbiol.*, 1991, **137**, 2811-2815.
55. M. Pohl, in *New Enzymes for Organic Synthesis: Screening, Supply and Engineering*, ed. T. Scheper, Springer, Germany, 1997, ch. 2. Protein Design on Pyruvate Decarboxylase (PDC) by Site-Directed Mutagenesis, pp. 15-41.
56. M. Pohl, G. A. Sprenger and M. Müller, *Curr. Opin. Biotechnol.*, 2004, **15**, 335-342.
57. M. B. Doughty and G. E. Risinger, *Bioorganic Chem.*, 1987, **15**, 1-14.
58. R. Silverman, in *Organic Chemistry of Enzyme-Catalyzed Reactions*, Academic Press, Cornwall, UK, 2002, ch. 8. Decarboxylation, pp. 321-357.
59. F. Jordan, *Nat. Prod. Rep.*, 2003, **20**, 184-201.
60. F. Dyda, W. Furey, S. Swaminathan, M. Sax, B. Farrenkopf and F. Jordan, *Biochemistry*, 1993, **32**, 6165-6170.
61. S. König, *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.*, 1998, **1385**, 271-286.
62. H. Zehender, D. Trescher and J. Ullrich, *Eur. J. Biochem.*, 1987, **167**, 149-154.
63. J. Rivoal, B. Ricard and A. Pradet, *Eur. J. Biochem.*, 1990, **194**, 791-797.
64. U. Mucke, S. Konig and G. Hubner, *Biol. Chem. Hoppe-Seyler*, 1995, **376**, 111-117.
65. Y. A. Muller, G. Schumacher, R. Rudolph and G. E. Schulz, *J. Mol. Biol.*, 1994, **237**, 315-335.
66. R. G. Duggleby, *Acc. Chem. Res.*, 2006, **39**, 550-557.
67. S. Kutter, M. S. Weiss, G. Wille, R. Golbik, M. Spinka and S. Konig, *J. Biol. Chem.*, 2009, **284**, 12136-12144.
68. C. F. Hawkins, A. Borges and R. N. Perham, *FEBS Lett.*, 1989, **255**, 77-82.
69. *USA Pat.,* US5723584A, 1998.
70. A. D. Gounaris, I. Turkenkopf, S. Buckwald and A. Young, *J. Biol. Chem.*, 1971, **246**, 1302-1309.
71. J. H. R. Kägi and B. L. Vallee, *J. Biol. Chem.*, 1960, **235**, 3188-3192.
72. Promega, Calculating Nucleic Acid or Protein Concentration: Technical Note).
73. H. Bergmeyer, in *Methods of Enzymatic Analysis*, eds. J. Bergmeyer and M. Graßl, VCH Verlagsgesellschaft mbH, Weinheim, 1983, vol. II, ch. 2. Pyruvate decarboxylase, pp. 302-303.
74. D. Gocke, PhD, University of Düsseldorf, 2007.
75. P. K. Agarwal, V. Uppada, A. G. Swaminathan and S. B. Noronha, *Bioresour. Technol.*, 2015, **192**, 90-96.
76. T. Matsuda, K. Nakayama, T. Abe and M. Mukouyama, *Biocatal. Biotransform.*, 2010, **28**, 167-171.
77. A. Boiteux and B. Hess, *FEBS letters*, 1970, **9**, 293-296.
78. S. Pal, D.-H. Park and B. V. Plapp, *Chem.-Biol. Interact.*, 2009, **178**, 16-23.

References

79.   G. S. Parks and H. M. Huffman, *Free Energies of Some Organic Compounds* Chemical Catalogue Co., New York, 1932.
80.   S. Carson, S. Ruben, M. Kamen and J. Foster, *Proc. Natl. Acad. Sci. U. S. A.*, 1941, **27**, 475.
81.   M. de la Plaza, P. F. de Palencia, C. Pelaez and T. Requena, *Fems Microbiology Letters*, 2004, **238**, 367-374.
82.   B. A. Smit, J. E. T. van Hylckama Vlieg, W. J. M. Engels, L. Meijer, J. T. M. Wouters and G. Smit, *Appl. Environ. Microbiol.*, 2005, **71**, 303.
83.   A. Yep, G. L. Kenyon and M. J. McLeish, *Bioorganic Chem.*, 2006, **34**, 325-336.
84.   C. L. Berthold, D. Gocke, M. D. Wood, F. J. Leeper, M. Pohl and G. Schneider, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2007, **63**, 1217-1224.
85.   Y. Lindqvist and G. Schneider, *Curr. Opin. Struct. Biol.*, 1993, **3**, 896-901.
86.   D. Gocke, C. L. Nguyen, M. Pohl, T. Stillger, L. Walter and M. Mueller, *Adv. Synth. Catal.*, 2007, **349**, 1425-1435.
87.   M. Höhne and U. T. Bornscheuer, in *Enzyme Catalysis in Organic Synthesis*, Wiley-VCH, Germany, 2012, vol. 1, ch. 19. Application of Transaminases, pp. 779-820.
88.   A. Braunshtein and M. Shemiakin, *Biochemistry*, 1953, **18**, 393.
89.   J. D. Lueck and H. J. Fromm, *FEBS Lett.*, 1973, **32**, 184-186.
90.   A. K. B. Bour, Doctor of Philosophy, UCL (University College London), 2016.
91.   W. Jenkins and M. L. Fonda, in *Transaminases*, eds. P. Christen and D. E. Metzler, John Wiley and Sons, New York, 1985, pp. 216-234.
92.   M. Thomsen, L. Skalden, G. J. Palm, M. Hohne, U. T. Bornscheuer and W. Hinrichs, *Acta Crystallographica Section F-Structural Biology and Crystallization Communications*, 2013, **69**, 1415-1417.
93.   P. K. Mehta, T. I. Hale and P. Christen, *Eur. J. Biochem.*, 1993, **214**, 549-561.
94.   J. Ward and R. Wohlgemuth, *Curr. Org. Chem.*, 2010, **14**, 1914-1927.
95.   W. Leuchtenberger, K. Huthmacher and K. Drauz, *Appl. Microbiol. Biotechnol.*, 2005, **69**, 1-8.
96.   E. Ricca, B. Brucher and J. H. Schrittwieser, *Adv. Synth. Catal.*, 2011, **353**, 2239-2262.
97.   A. Galkin, L. Kulakova, T. Yoshimura, K. Soda and N. Esaki, *Appl. Environ. Microbiol.*, 1997, **63**, 4651-4656.
98.   A. Menzel, H. Werner, J. Altenbuchner and H. Gröger, *Eng. Life Sci.*, 2004, **4**, 573-576.
99.   H. Gröger, O. May, H. Werner, A. Menzel and J. Altenbuchner, *Org. Process Res. Dev.*, 2006, **10**, 666-669.
100.   P. P. Taylor, D. P. Pantaleone, R. F. Senkpeil and I. G. Fotheringham, *Trends Biotechnol.*, 1998, **16**, 412-418.
101.   H. Ziehr, M. R. Kula, E. Schmidt, C. Wandrey and J. Klein, *Biotechnol. Bioeng.*, 1987, **29**, 482-487.
102.   Y. Hu, T. Tang, W. Yang and H. Zhou, *Process Biochem.*, 2009, **44**, 142-145.
103.   T. Li, A. B. Kootstra and I. G. Fotheringham, *Org. Process Res. Dev.*, 2002, **6**, 533-538.
104.   L. Cerioli, M. Planchestainer, J. Cassidy, D. Tessaro and F. Paradisi, *J. Mol. Catal. B: Enzym.*, 2015, **120**, 141-150.
105.   J. S. Shin and B. G. Kim, *Biotechnol. Lett.*, 2009, **31**, 1595-1599.
106.   A. P. Green, N. J. Turner and E. O'Reilly, *Angew. Chem., Int. Ed.*, 2014, **53**, 10714-10717.
107.   U. Schell, R. Wohlgemuth and J. M. Ward, *J. Mol. Catal. B: Enzym.*, 2009, **59**, 279-285.
108.   P. Felig, *Metab., Clin. Exp.*, 1973, **22**, 179-207.
109.   F. Garcia-Campusano, V.-H. Anaya, L. Robledo-Arratia, H. Quezada, H. Hernandez, L. Riego and A. Gonzalez, *Can. J. Microbiol.*, 2009, **55**, 368-374.
110.   G. Peñalosa-Ruiz, C. Aranda, L. Ongay-Larios, M. Colon, H. Quezada and A. Gonzalez, *PLoS One*, 2012, **7**, e45702.
111.   E. Rojas-Ortega, B. Aguirre-Lopez, H. Reyes-Vivas, M. Gonzalez-Andrade, J. C. Campero-Basaldua, J. P. Pardo and A. Gonzalez, *Front. Microbiol.*, 2018, **9**, 1-16.
112.   D. M. Francis and R. Page, *Curr. Protoc. Protein Sci.*, 2010, **61**, 1-29.
113.   N. Regev-Rudzki, O. Yogev and O. Pines, *J. Cell Sci.*, 2008, **121**, 2423-2431.

References

114. E. P. A. Neve and M. Ingelman-Sundberg, *J. Biol. Chem.*, 2001, **276**, 11317-11322.
115. I. Umemura, K. Yanagiya, S. Komatsubara, T. Sato and T. Tosa, *Biosci., Biotechnol., Biochem.*, 1994, **58**, 283-287.
116. K. Zhang, H. Li, K. M. Cho and J. C. Liao, *Proc. Natl. Acad. Sci.*, 2010, **107**, 6234-6239.
117. R. Tao, Y. Jiang, F. Zhu and S. Yang, *Biotechnol. Lett.*, 2014, **36**, 835-841.
118. R. M. Aguiar, R. A. Leão, A. Mata, D. Cantillo, C. O. Kappe, L. S. Miranda and R. O. de Souza, *Org. Biomol. Chem.*, 2019, **17**, 1552-1557.
119. V. Raju, S. Somaiah, S. Sashikanth, E. Laxminarayana and K. Mukkanti, *Indian J. Chem.*, 2014, **53B**, 1218-1221.
120. Y. Perez-Fuertes, J. E. Taylor, D. A. Tickell, M. F. Mahon, S. D. Bull and T. D. James, *J. Org. Chem.*, 2011, **76**, 6038-6047.
121. P. Anastas and N. Eghbali, *Chem. Soc. Rev.*, 2010, **39**, 301-312.
122. N. S. Punekar, in *ENZYMES: Catalysis, Kinetics and Mechanisms*, Springer, Singapore, 1 edn., 2018, ch. 4. Chemical Mechanisms and Catalysis, Carboxylations and Decarboxylations, pp. 403-420.
123. H. R. Hobbs and N. R. Thomas, *Chem. Rev.*, 2007, **107**, 2786-2820.
124. V. V. Mozhaev, K. Heremans, J. Frank, P. Masson and C. Balny, *Proteins: Struct., Funct., Bioinf.*, 1996, **24**, 81-91.
125. B. Nozière, P. Dziedzic and A. Córdova, *Phys. Chem. Chem. Phys.*, 2010, **12**, 3864-3872.
126. B. Nozière and P. Chabert, *Int. J. Chem. Kinet.*, 2010, **42**, 676-686.
127. P. K. Glasoe and F. Long, *J. Phys. Chem.*, 1960, **64**, 188-190.
128. H. J. Buschmann, H. H. Füldner and W. Knoche, *Ber. Bunsen. Phys. Chem.*, 1980, **84**, 41-44.
129. M. G. Kissane, S. A. Frank, G. A. Rener, C. P. Ley, C. A. Alt, P. A. Stroud, R. K. Vaid, S. K. Boini, L. A. McKee and J. T. Vicenzi, *Tetrahedron Lett.*, 2013, **54**, 6587-6591.
130. D. P. Kjell, B. J. Slattery and M. J. Semo, *J. Org. Chem.*, 1999, **64**, 5722-5724.
131. *IT Pat.,* EP0066113, 1986.
132. H. Cai, T. J. Mangner, O. Muzik, M.-W. Wang, D. C. Chugani and H. T. Chugani, *ACS Med. Chem. Lett.*, 2014, **5**, 1152-1155.
133. M. Planchestainer, D. Roura Padrosa, M. L. Contente and F. Paradisi, *Catalysts*, 2018, **8**, 40.
134. M. Planchestainer, M. L. Contente, J. Cassidy, F. Molinari, L. Tamborini and F. Paradisi, *Green Chem.*, 2017, **19**, 372-375.
135. M. T. Reetz, J. D. Carballeira, J. Peyralans, H. Höbenreich, A. Maichele and A. Vogel, *Chem. Eur. J.*, 2006, **12**, 6031-6038.
136. U. T. Bornscheuer and M. Hohne, in *Protein Engineering*, Humana Press, New Jersey 1edn., 2018, ch. 1. Protein Engineering: Past, Present, and Future, pp. 1-12.
137. C. Fagain, *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.*, 1995, **1252**, 1-14.
138. U. Hahn and U. Heinemann, in *Concepts in Protein engineering and Design*, eds. P. Wrede and G. Schneider, Walter de Gruyter, Berlin, Germany, 1994, ch. 3. Structure determination, modeling and site-directed mutagenesis studies, p. 109.
139. A. Fresht and G. Winter, *Trends Biochem. Sci.*, 1986, **17**, 292-294.
140. NEB, Q5® Site-Directed Mutagenesis Kit, https://www.neb.com/products/e0554-q5-site-directed-mutagenesis-kit., (accessed 21 Januaray, 2021).
141. J. M. Candy and R. G. Duggleby, *Biochim. Biophys. Acta*, 1998, **1385**, 323-338.
142. X. Y. Pei, K. M. Erixon, B. F. Luisi and F. J. Leeper, *Biochemistry*, 2010, **49**, 1727-1736.
143. G. Lu, D. Dobritzsch, S. Baumann, G. Schneider and S. König, *Eur. J. Biochem.*, 2000, **267**, 861-868.
144. F. Dyda, W. Furey, S. Swaminathan, M. Sax, B. Farrenkopf and F. Jordan, *Biochem. J.*, 1993, **32**, 6165-6170.
145. M. Pohl, P. Siegert, K. Mesch, H. Bruhn and J. Grotzinger, *Eur. J. Biochem.*, 1998, **257**, 538-546.
146. F. H. Andrews and M. J. McLeish, *Bioorganic Chem.*, 2012, **43**, 26-36.
147. M. Liu, E. A. Sergienko, F. Guo, J. Wang, K. Tittmann, G. Hübner, W. Furey and F. Jordan, *Biochemistry*, 2001, **40**, 7355-7368.
148. A. K. Chang, P. F. Nixon and R. G. Duggleby, *Biochem. J.*, 1999, **339**, 255-260.

149. C. Y. Huang, A. K. Chang, P. F. Nixon and R. G. Duggleby, *Eur. J. Biochem.*, 2001, **268**, 3558-3565.
150. E. A. Sergienko and F. Jordan, *Biochemistry*, 2001, **40**, 7369-7381.
151. Y. G. Wu, A. K. Chang, P. F. Nixon, W. Li and R. G. Duggleby, *Eur. J. Biochem.*, 2000, **267**, 6493-6500.
152. F. Jordan, I. Baburina, Y. Gao, F. Guo, A. Kahyaoglu, N. Nemeria, A. Volkov, J. Yi, D. Zhang and R. Machado, in *Biochemistry and physiology of thiamin diphosphate enzymes*, eds. H. Bisswanger and A. Schellenberger, A. u. C. Intemann Verlag, Prien, Germany, 1996, ch. New insights to the regulation of thiamin diphosphate dependent decarboxylases by substrate and THDP. Mg (II), pp. 53-69.
153. D. Gocke, T. Graf, H. Brosi, I. Frindi-Wosch, L. Walter, M. Muller and M. Pohl, *Journal of Molecular Catalysis B-Enzymatic*, 2009, **61**, 30-35.
154. H. Iding, P. Siegert, K. Mesch and M. Pohl, *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.*, 1998, **1385**, 307-322.
155. P. Siegert, PhD, Heinrich-Heine University Duesseldorf, 2000.
156. H. Holzer, G. Schultz, C. Villar-Palasi and J. Juntgen-Sell, *Biochemische Zeitschrift*, 1956, **327**, 331-344.
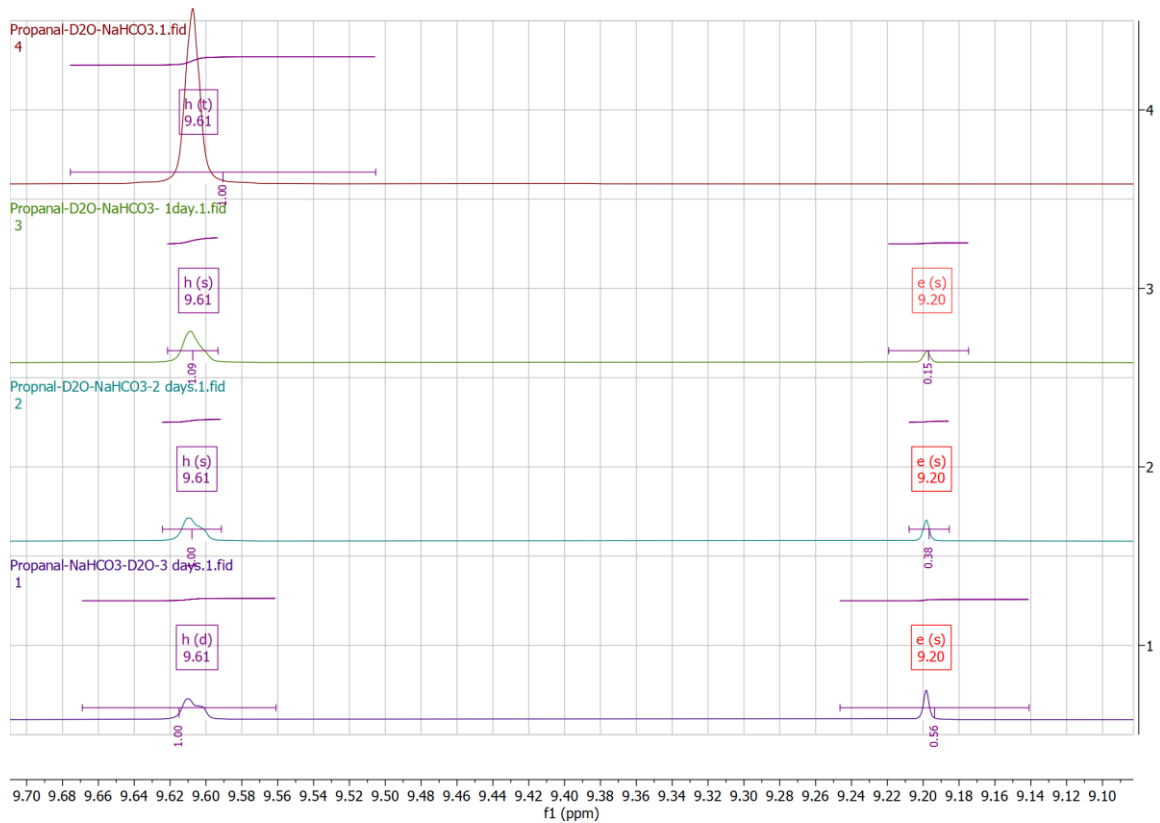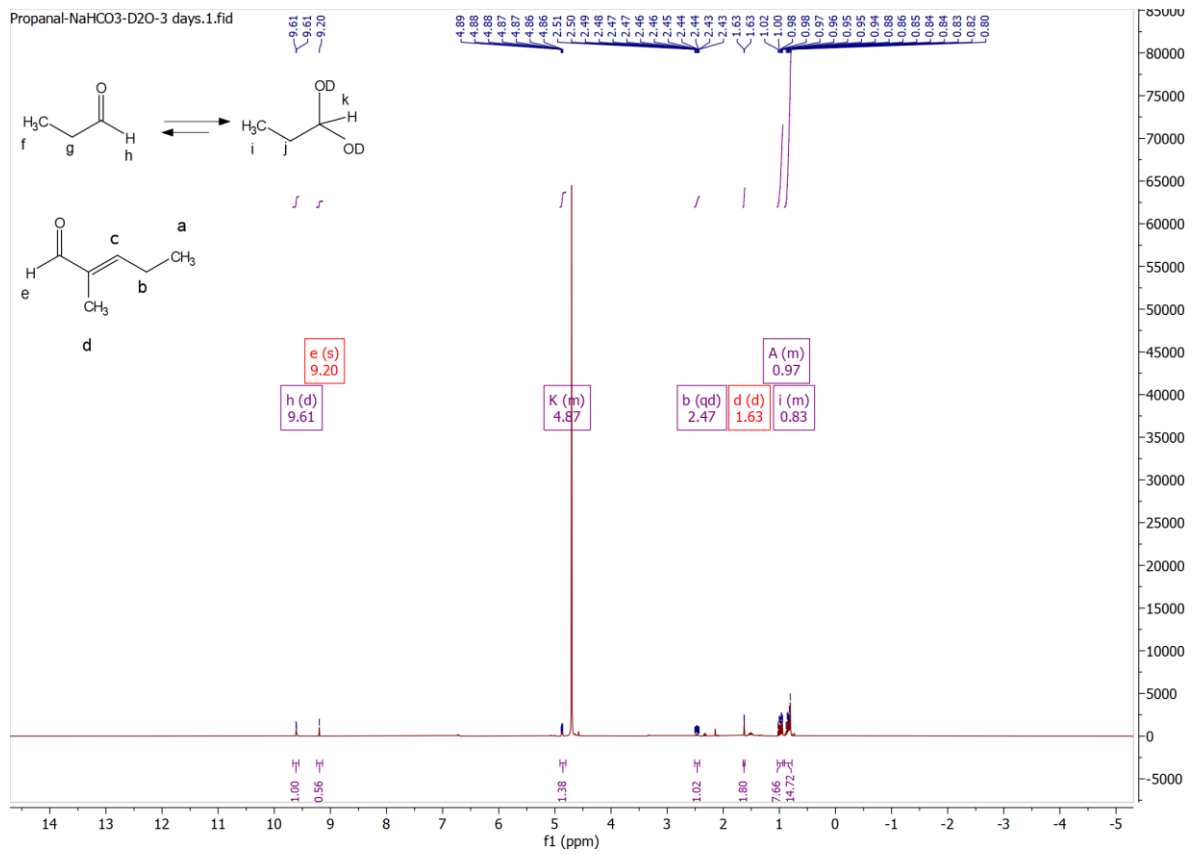
Appendix

A.1. ^{1}H NMR, 400 MHz

Investigation of the stability of the propanal for 3 days in carbonate buffer
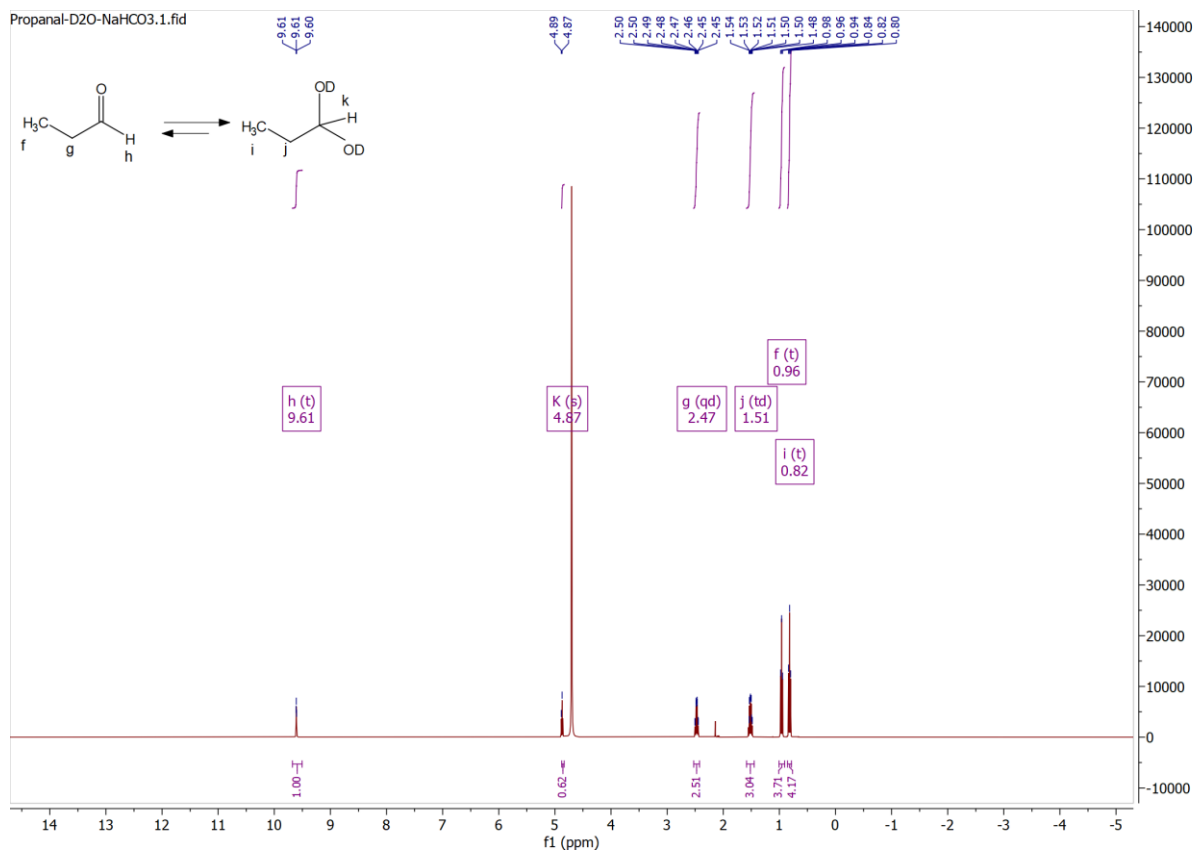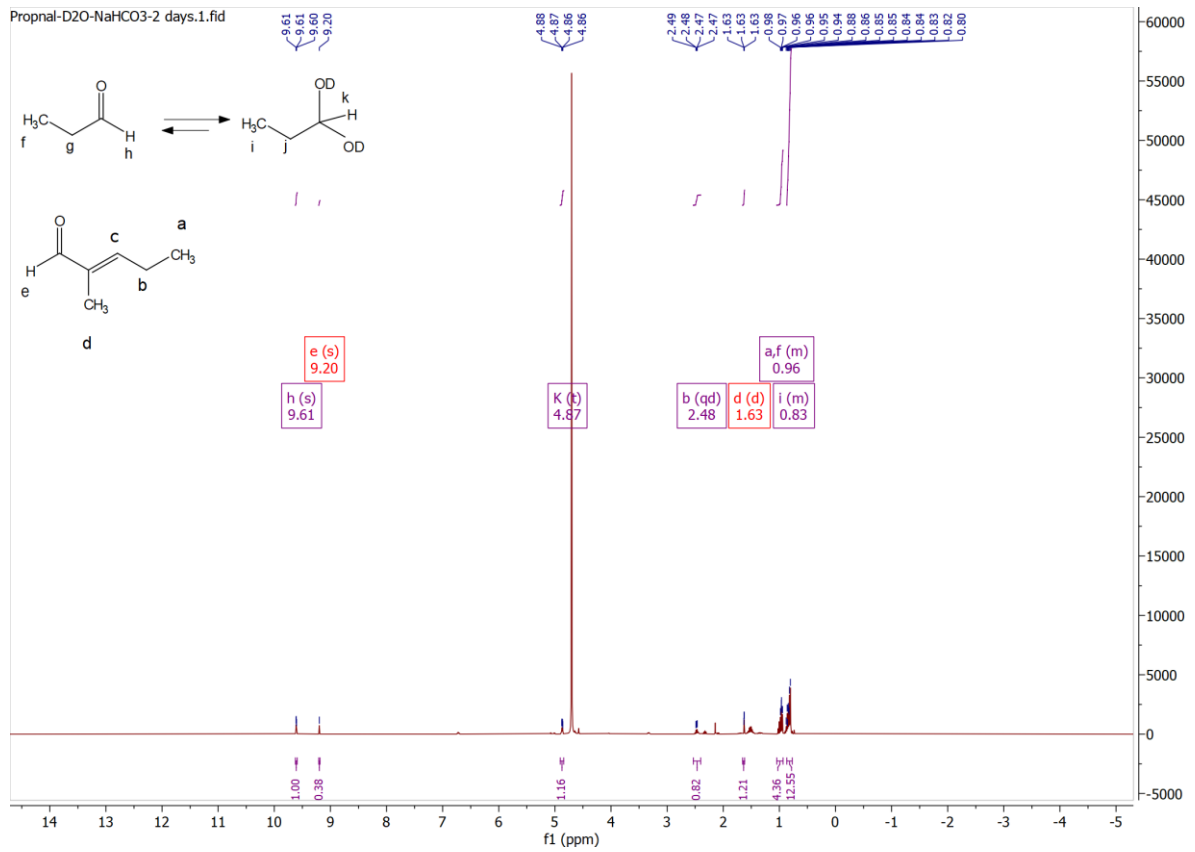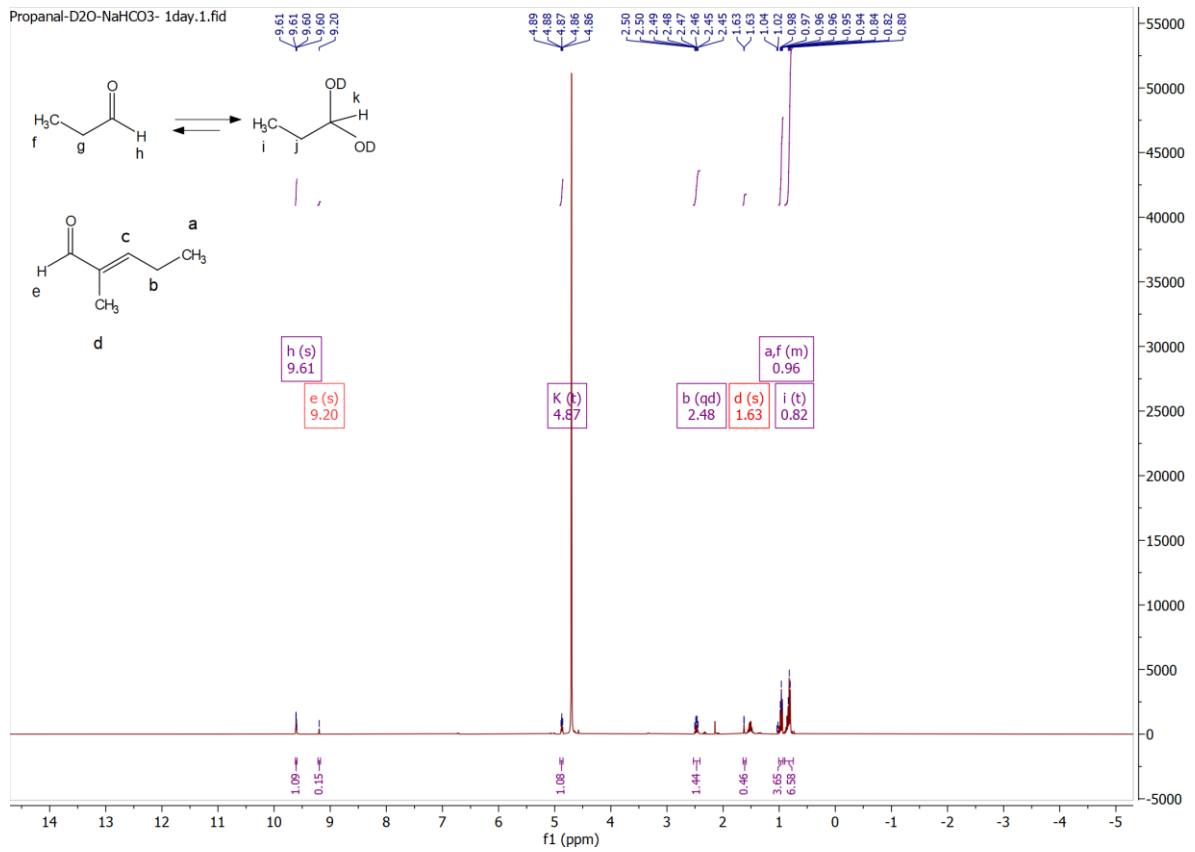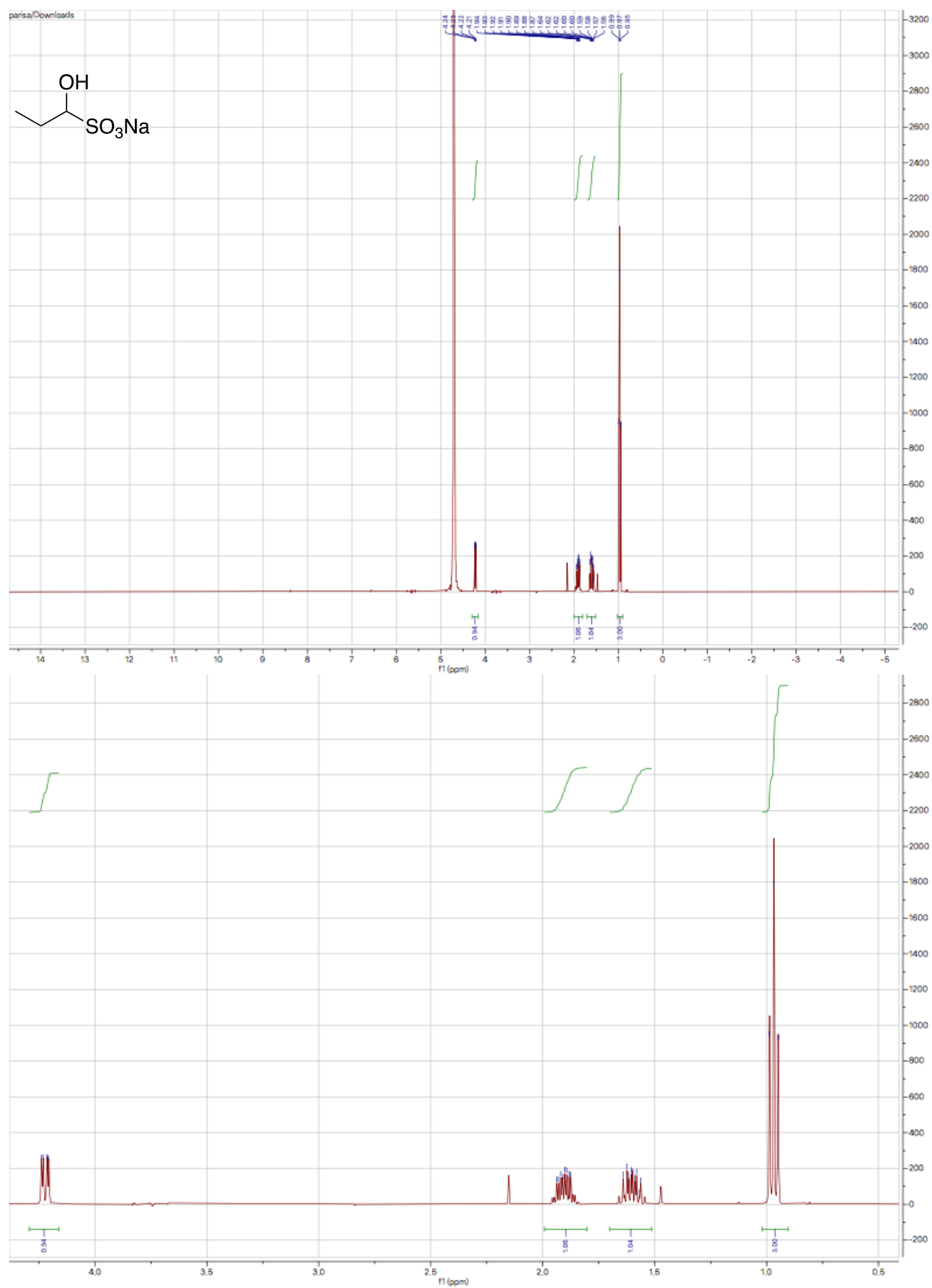
Preparation of sodium bisulfite

## A.2. Mass spectrum

Mass spectrum of the propanal bisulfite

Mass Spectrometry
Analytical Services
School of Chemistry

The University of Nottingham

Sample-ID	m_lam_Propanal bisulfite pH h2o	Lab	C13
Submitter	Malcolm Lamont	Supervisor	Anca Pordea
Analysis Name	m_lam_Propanal bisulfite pH h2o_24_01_92771.d	Acquisition Date	11/12/2020 4:22:39 PM
Ionisation Mode	ESI – Positive	Instrument	Bruker MicroTOF

+MS, 0.7-0.9min #59-75
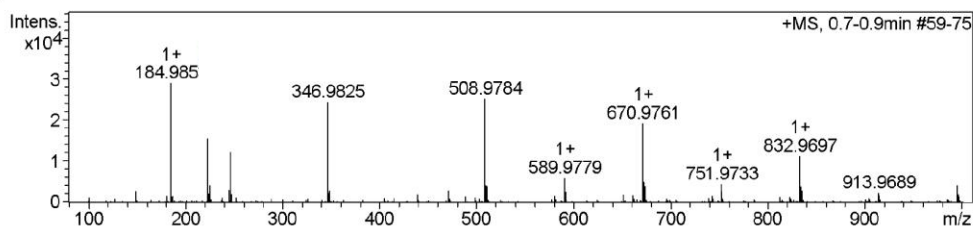
| #	m/z	I %
1	184.9853	100.0
2	223.1437	53.3
3	225.1953	14.0
4	245.1257	10.0
5	247.1782	42.3
6	346.9825	83.9
7	348.9799	9.5
8	471.3670	9.5
9	508.9784	86.5
10	509.9806	14.1
11	510.9755	13.4
12	589.9779	20.6

A.3. HPLC Chromatogram

HPLC calibration curve for L-homoalanine pre-derivatised with FMOC-Cl. Error bars in red indicate SD for three repeat experiments.
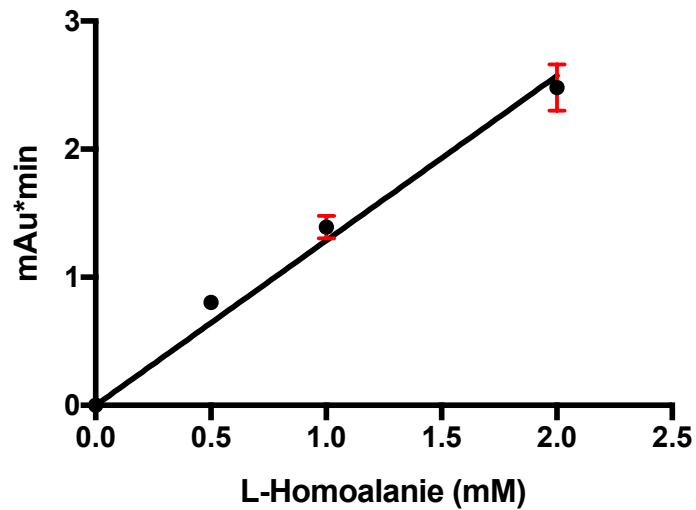
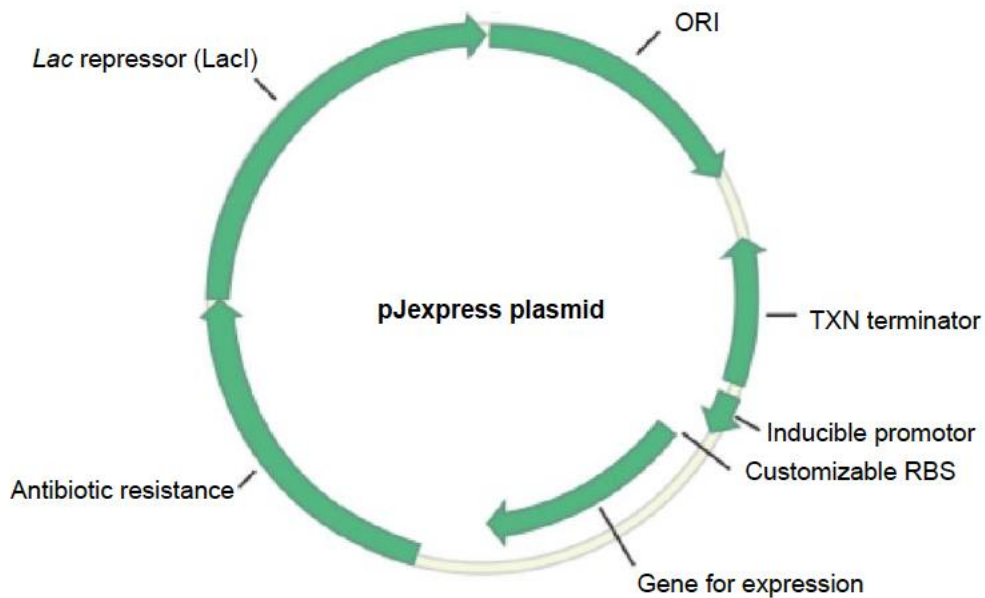HPLC Chromatogram of carboxylation-transamination reaction monitored at different intervals.
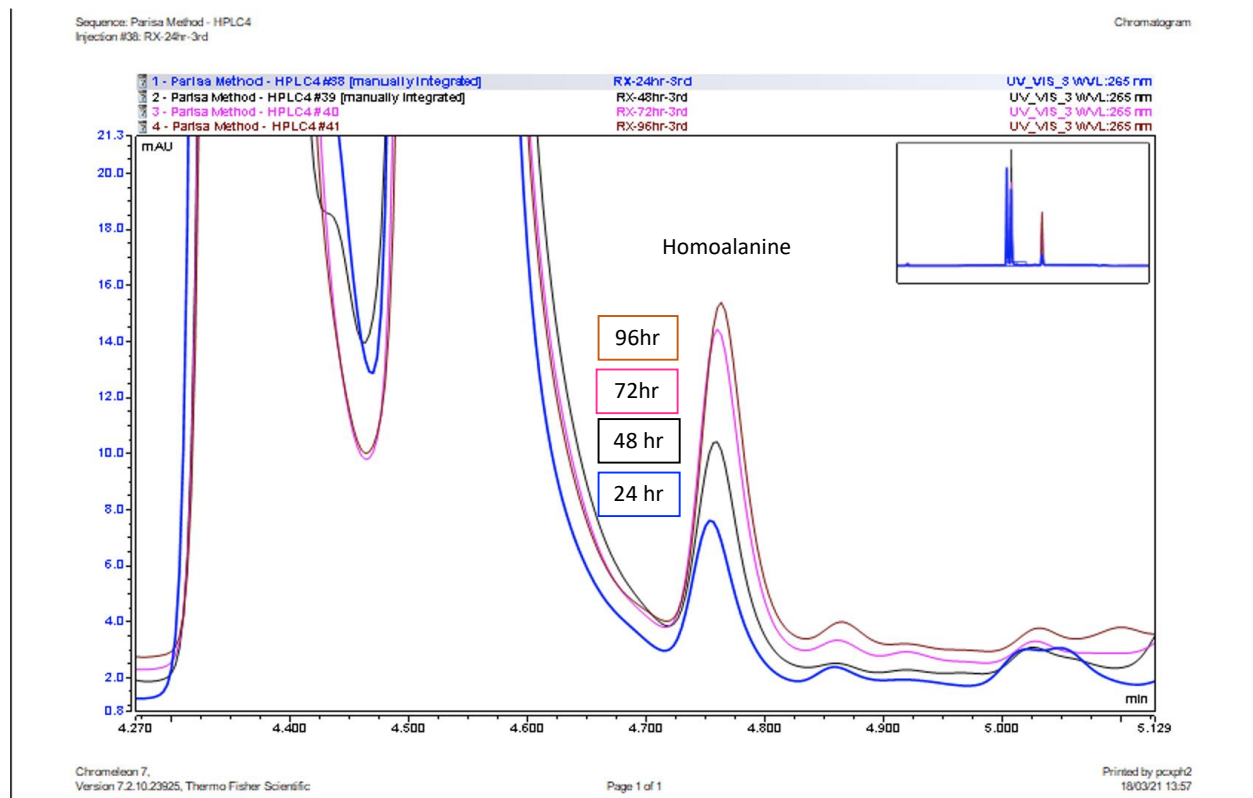
HPLC Chromatogram of carboxylation-transamination reaction monitored at different pH
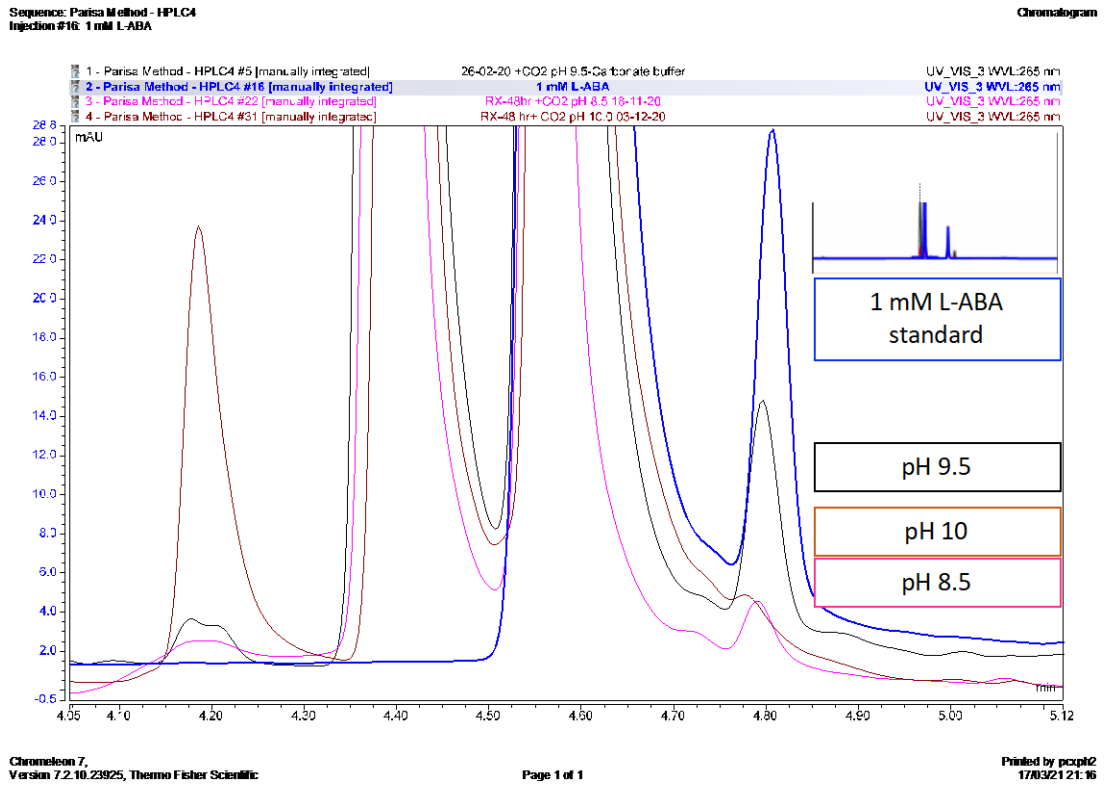
A.4. Plasmid Map

The plasmid map of pJexpress414 was given kindly by Dr. Lei Zhang. This plasmid was ordered originally from DNA 2.0 (USA).

The plasmid map of pET32a-*Ll*KdcA as ordered from DC Bioscience Ltd (UK).