ENGINEERING *S. CEREVISIAE* FOR ALKANE BIOSYNTHESIS FROM ALDEHYDE DECARBONYLASE PRECURSOR

HAZARKI YAOHARI

Division of Bioengineering

School of Chemical and Biomedical Engineering

A thesis submitted to the Nanyang Technological University in partial fulfillment of the requirement for the degree of Master of Engineering

2013

Abstract

Aldehyde decarbonylase (AD) catalyzes the conversion of aldehyde to alkane. ADs are naturally present in plants, insect or cyanobacteria. Cyanobacterial aldehyde decarbonylase (cAD) is a member of ferritin-like di-iron enzymes that catalyzes the conversion of aldehydes to alkanes and formate in the presence of a reducing system. This cAD enzyme can be utilized to generate a biofuel-producing platform by engineering biological systems to exploit this pathway for alkane synthesis from fatty acyl precursors. This study reports the expression, purification and characterization of two cADs that have not been studied, i.e. cADs from *Anabaena variabilis* ATCC 29413 (AD1) and *Arthrospira platensis* str. Paraca (AD2). Both genes which were cloned into *Saccharomyces cerevisiae* exhibit soluble expression and were subsequently extracted in its native form using a mild detergent, Tween 20, for *in vitro* characterization. AD1 showed activity towards dodecanal substrate with K_M = 3.5±2.9 µM and k_{cat} = 1.4±0.1/h under the phenazine methosulfate (PMS)/NADH reducing system. AD2 has a K_M and k_{cat} of 5.0±2.2 µM and 0.9±0.1/min, respectively. When NADH was replaced with NADPH, an electron donor that is abundant inside the cell, the efficiency of the cADs towards dodecanal substrate was compromised, though NADPH can still support the reaction.

The activity of a few other decarbonylases from cyanobacteria and insect *in vivo* were also investigated. The ADs from *Synechococcus elongatus* (AD3) and *Musca domestica* (cytochrome P450 decarbonylase CYP4G2-CPR) were successfully cloned into *S. cerevisiae* and recombinantly expressed. I*n vivo* conversion of a range of exogenous aldehydes fed to the cells expressing these decarbonylases did not yield alkanes. Further investigations

involving protein and pathway engineering to improve decarbonylase activity are underway to address this challenge.

Acknowledgements

I would like to thank Nanyang Technological University (NTU) for the opportunity to do this Master of Engineering (M. Eng.) Programme in the School of Chemical and Biomedical Engineering. I would also like to thank the National Research Foundation for funding this project through the Competitive Research Programme. I would like to express my gratitude towards Professor Susanna Leong Su Jan and Professor Matthew Wook Chang, my supervisors, my lab members: Dr. Foo Jee Loon, Dr. Sun Zhoutong, Dr. Ling Hua, Dr. Zhao Hongxin and Lim Pei Yu of Nanyang Technological University who have given their help and support. I welcome any suggestions and/or criticisms from interested readers.

Hazarki Yaohari

August 2013

Table of Contents

List of Figures

1. Introduction

An alkane/alkene biosynthesis pathway was recently discovered in cyanobacteria [1]. The pathway consists of two sequential reactions which involve two main enzymes, a fatty acyl-CoA reductase (FAR) and an aldehyde decarbonylase (AD) (Figure 1). FAR reduces fatty acid intermediates (fatty acyl-CoA or fatty acyl-ACP) into fatty aldehydes, while AD converts fatty aldehydes into corresponding alkanes or alkenes. This novel pathway can be utilized to produce alkanes, a major component in gasoline, to replace non-renewable resources. This outcome opens the way for scientist to generate a renewable and low-cost alkane-producing platform by engineering biological systems to exploit this pathway to synthesize alkanes from fatty acids.

The objective of this study is to engineer *S. cerevisiae* for C8-C18 alkane biosynthesis, by heterologous co-expression of gene candidates encoding FAR and AD. Specifically, my research project aims to achieve the following objectives:

(i) Characterization and optimization of the aldehyde decarbonylase enzymatic reaction *in vitro*.

(ii) *In vivo* alkane production in *S. cerevisiae* using engineered pathway.

(iii) Protein engineering to improve AD specificity and activity towards certain aldehyde substrates to increase alkane production yield and throughput.

(iv) Pathway construction *in vivo* optimization to integrate AD with FAR for conversion of fatty acids to alkanes in *S. cerevisiae*.

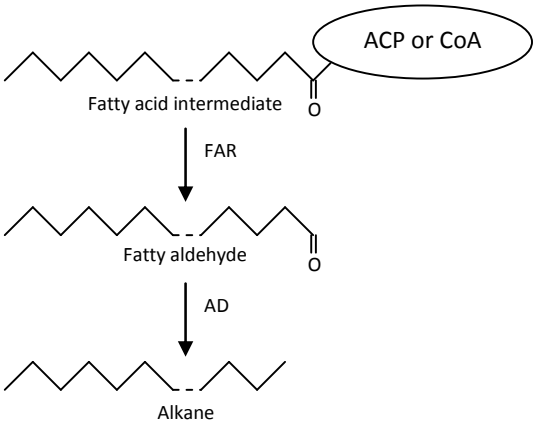


Figure 1 Schematic of the alkane biosynthesis pathway found in cyanobacteria involving a fatty acyl-ACP/CoA reductase and an aldehyde decarbonylase.

2. Literature Review

The aldehyde decarbonylase pathway, as well as the selection of *S. cerevisiae* as microbial host and its fatty acid metabolism pathway will be discussed in this section. These are important factors that might affect the production of alkanes through the recombinant AD pathway.

2.1 Aldehyde decarbonylase

Alkanes, the major component in gasoline, diesel and jet fuel, are naturally synthesized by varied organisms, for example as cuticular waxes in plants [2-5], as cuticular lipids [6] or sex pheromones [7] in insects and in cyanobacteria [1, 8] as recently discovered. Studies revealed that genes encoding aldehyde decarbonylases are responsible for the production of alkanes in these organisms. According to Eser et al. [9], there are several classes of ADs found in different organisms. The first class, found primarily in plants, is a metal-dependent integral membrane protein that produces alkanes and carbon monoxide (CO) from aldehydes [10, 11]. ADs in higher plants are mostly expressed from orthologs of *cer*1 of *Arabidopsis thaliana*, which appears to be conserved throughout many species [2]. The second, found in insects, is a cytochrome P450 which forms carbon dioxide (CO_2) instead of CO. The third is cyanobacterial AD (cAD) which is a relatively small and soluble protein related to the non-heme di-iron reductases [1]. Although plant decarbonylases have been identified for a long time and are conserved among diverse species [2, 3, 12, 13], no studies successfully isolate and characterize the exact functions of the enzymes. ADs from insects, a cytochrome P450, is a large enzyme complex and therefore, more difficult to manipulate.

Hence, of the three classes, cyanobacterial ADs show the most promising potential to be utilized as an alkane-producing platform.

2.1.1 Plant decarbonylase

A number of gene products from plants and cyanobacetria have been reported to have the decarbonylase activity. Hannoufa *et al*. [3] showed that there is a considerable accumulation of aldehydes and reduction of alkanes in the epicuticular wax of *Arabidopsis thaliana cer*1 mutant, indicating that *cer*1 gene may encode an AD. From wax chemical composition analysis, it was known that the conversion of C30 aldehyde to C29 alkane is blocked. Similarly, a putative AD (*Hvcer1;1*) with sequence homology to the *cer*1 in was isolated from Hordeum vulgare [12]. These plant decarbonylases contain three histidine-rich motifs with eight conserved histidine residues that are typical for iron binding proteins.

2.1.2 Insect decarbonylase

In insects, the conversion of very long chain aldehydes (C21-37+) to hydrocarbons is catalyzed by an aldehyde oxidative decarbonylase P450 (CYP). This CYP requires another enzyme to act as a redox partner, NADPH-cytochrome P450 reductase (CPR), for oxidative decarbonylastion of aldehyde. In *Drosophila melanogaster*, CYP4G1 and CPR enzymes are highly coexpressed and colocalized into oenocytes in a near optimal stoichiometric ratio [14, 15]. These oenocytes are specialized cells in the abdominal integument of insects in which hydrocarbons are synthesized [16].

The CYP4G subfamily of P450 was recently identified as aldehyde oxidative decarbonylase [14]. Repression of either CYP4G1 or CPR in *D. Melanogaster* resulted in high mortality rate of adult flies and reduced cuticular alkane/alkene level in surviving flies. The decrease of hydrocarbons in surviving CYP4G1-repressed flies caused significant decrease of desiccation tolerance and affected courtship behaviour of flies, as hydrocarbons are major components

in insect sex pheromones. Recombinant expression of a fusion of orthologous CYP4G2 (G2) from *Musca domestica* in Sf9 (insect) cells using baculovirus also showed NADPH-dependent oxidative decarbonylase activity on octadecanal. This fusion G2-CPR decarbonylase released carbon dioxide as a co-product (Figure 2), consistent with the proposed mechanism of native P450 enzymes [17].

$$R - CH_2 - CHO + O_2 + NADPH \xrightarrow{CYP4G+CPR} R - CH_2 - H + H_2O + NADP^+ + CO_2$$

Figure 2 Schemes representing reaction pathways by which aldehyde substrates undergo decarbonylation into alkanes by insect aldehyde oxidative decarbonylase, with carbon dioxide released as a co-product.

2.1.3 Cyanobacterial decarbonylase

Schirmer *et al.* [1] has described the successful construction of an alkane biosynthesis pathway using enzymes from the cyanobacteria *Synechococcus elongatus* in *E. coli*, which leads to the production and secretion of a range of alkanes of varying carbon chain lengths (C13 to C17). They observed that all alkanes produced were of odd carbon number, consistent with the "n-1" rule of decarbonylation of normally even-numbered aldehydes into alkanes [2, 11]. They suggested that cADs belong to the non-heme di-iron ribonucleotide reductase family. The crystal structure of a cAD from *Prochlorococcus marinus* MIT9313 shows resemblance to the second subunit (R2) of *E. coli* ribonucleotide reductase (Figure 3). Both proteins' active sites have di-iron centres coordinated to ligands (histidine residues).

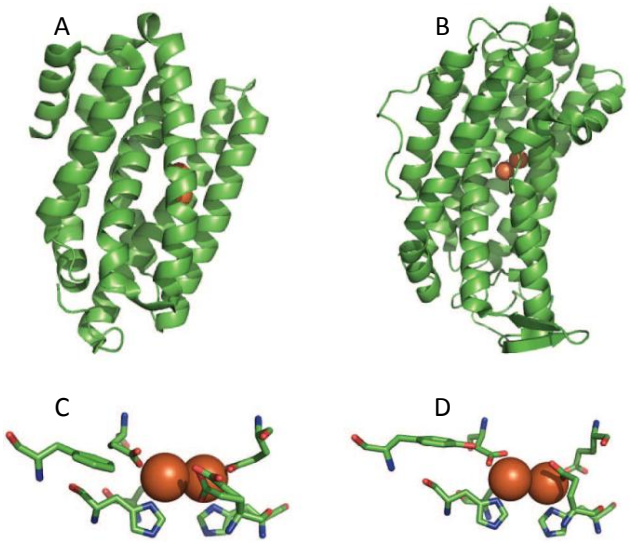


Figure 3 Three-dimensional models of (A) cAD from *P. marinus* MIT9313 (PDB ID 2OC5) and (B) Ribonucleotide reductase R2 from *E. coli* (PDB ID 1RIB), and active sites of (C) MIT9313 cAD and (D) *E. coli* R2; indicating di-iron centres and ligands (histidine residues).

Studies showed that some non-heme di-iron reductases require a reducing environment for activities. The reductase of a di-iron methane monooxygenase complex from *Methylococcus capsulatus* (Bath), for instance, contains two domains: an NADH-binding domain and a ferredoxin (Fd) domain [18]. The Fd domain carries a [2Fe-2S] cofactor and has an important role in electron-transfer to the di-iron centres. Another example is a [2Fe-2S] ferredoxin (YfaE) that is involved in the reduction of nucleotides to deoxynucleotides by an *E. coli* ribonucleotide reductase (RNR). The YfaE is thought to be involved in the maintenance and biosynthetic pathways of a diferric-tyrosyl radical cofactor in the β subunit of the RNR that is crucial for the reduction [19].

Indeed, cADs are iron-dependent and an external reducing system is necessary for activity. Decarbonylation of aldehydes is only observed when cofactors-which consist of reducing

equivalents such as spinach ferredoxin, ferredoxin reductase and NADPH-are present [1]. When Fe(II) was removed from the active metal centre of the enzyme and substituted by other divalent metals such as Mn(II), Cu(II), Zn(II), Co(II) and Ni(II), decarbonylase activity was significantly reduced compared to that obtained when Fe(II) were present [8].

In the decarbonylation reaction, the aldehyde substrate is hydrolyzed into an alkane, forming a formate group as a stoichiometric co-product [8, 20] (Figure 4B). In addition, Das *et al.* (2011) discovered that using phenazine methosulfate (PMS):NADH reducing system instead of ferredoxin system can greatly improve enzyme activity [8]. PMS was thought to interact with and bind to the cAD, which is critical for electron transports during the reaction, while NADH is not consumed by the reaction and can be replaced by other reductants. This (PMS):NADH system are often used in enzymatic assays as a reducing substrate. PMS serves as an electron mediator between NADH and oxidizing cofactors while NADH is a two-electron donor [21-24]. They also hypothesized that bovine serum albumin (BSA) might interact with the enzyme, playing an important role in delivering the insoluble aldehyde substrate to the enzyme [9]. BSA has been known to bind long-chain fatty acids [25] and other cationic lipids [26], increasing their solubility in aqueous solutions. The binding of BSA to these cationic lipids occurs via hydrophilic and hydrophobic contacts.

It is unclear, however, whether the decarbonylation is an oxygen-dependent reaction as contradictory results were reported independently. Li *et al.* [27] reported that cAD was oxygen-dependent while Das *et al.* found that cAD does not require oxygen and in fact, acivity was 100-fold higher in the anaerobic condition [8, 9]. The reason behind this observation was later explained by the group to be due to some inconsistencies [28].

Due to the discrepancy in reaction condition, two fundamentally different decarbonylation pathways were proposed (Figure 4): a cryptic oxidation of aldehyde substrate by O_2 [27] and an anaerobic hydrolysis of aldehydes to alkanes [8, 9].

A $$R - CH_2 - CHO + 4[H^+] + 4[e^-] + O_2^* \xrightarrow{cAD} R - CH_3 + HCO^*OH + H_2O^*$$

B $$R - CH_2 - CHO + H_2^*O \xrightarrow{cAD} R - CH_2 - H^* + HCO - OH^*$$

Figure 4 Schemes representing reaction pathways by which aldehyde substrates undergo decarbonylation into alkanes. (A) cryptic oxidation pathway and (B) oxygen-independent hydrolysis of aldehydes. Both reactions results in the same products, an alkane and a formate.

A more recent study suggested that the enzymatic conversion is, in fact, a cryptic redox oxygenation process that strictly requires O_2, with formate co-product, instead of carbon monoxide [29]. Thus, the enzyme was proposed to be redesignated as aldehyde-deformylating oxygenase (ADO). The requirement of molecular oxygen in the catalysis was confirmed [30] when an anaerobic ADO reaction setup in the presence of O_2-free argon showed no activity after incubation with octadecanal and redox cofactors.

Schirmer *et al*. [1] reported ten cyanobacteria with fully sequenced genomes that naturally produced alkanes. They identified one cAD, PCC7942_1593 from *S. elongatus*, which conferred alkane biosynthesis to *E. coli* through heterologous expression. They also classified a number of genes, with protein sequence similarities to PCC7942_1593, from different cyanobacteria as putative decarbonylases. Some of these genes' products have been tested for decarbonylase activity (Table 1).

Table 1 Selected cyanobacteria with putative AD genes

Cyanobacterium	% Protein sequence similarity (to PCC7942_1593)	Alkane production reported	Activity tested
Synechococcus elongatus PCC7942	100	Yes	Yes
Synechococcus elongatus PCC6301	100	Yes	Not tested
Arthrospira platensis str. Paraca	78	Not reported	Not tested
Nostoc sp. PCC7120	74	Yes	Yes
Anabaena variabilis ATCC29413	74	Yes	Not tested
Synechocystis sp. PCC6803	72	Yes	Yes
Cyanothece sp. ATCC51142	73	Yes	Yes
Nostoc punctiforme PCC73102	72	Yes	Yes
Acaryochloris marina MBIC11017	71	Not reported	Yes
Cyanothece sp. PCC7425	71	Yes	Yes
Thermosynechococcus elongatus BP-1	70	Not reported	Yes
Synechococcus sp. JA-3-3Ab	68	Not reported	Yes
Gloeobacter violaceus PCC7421	63	Yes	Yes
Synechococcus sp. RS9917	68	Not reported	Yes
Prochlorococcus marinus str. MIT9313	63	Not reported	Yes
Prochlorococcus marinus subsp. pastoris str. CCMP1986	60	Yes	Yes
Prochlorococcus marinus str. NATL2A	59	Not reported	Yes

2.2 The microbial host: *Saccharomyces cerevisiae*

Although functional cADs have been successfully expressed in *E. coli* as discussed above, the yeast *Saccharomyces cerevisiae* was chosen as the host in this study because it has been extensively studied and has many advantages [31] in the expression of clinically and industrially important proteins. Industrial scale yeast fermentation technologies are also widely available and can be used eventually, for the scale up of this study. Even though *E.*

coli is a prokaryote that can grow at a very fast rate and are able to express large amount of proteins in a fairly short time, they lack the ability to do post-translational protein modifications such as N- and O- glycosylation. Most of the time, they are also unable to produce correctly folded complex proteins, resulting in insoluble and inactive protein forms. There are known methods to assist protein folding *in-vivo* and improve its solubility. One way is to express the target protein as a fusion with soluble partners such as glutathione S-transferase (GST) or maltose-binding protein (MBP) [32]. However, expressing these large fusion proteins will exert a great metabolic burden to the cells. Yeast, on the other hand, is able to perform post-translational modifications and provide advanced protein folding pathway. Therefore, correctly folded and fully functional heterologous proteins can be produced without imposing additional stresses to the cells. Being a microorganism, yeast is able to grow much faster than more complex eukaryotes like mammalian cells. Unlike mammalian cells, yeast can grow on simple and relatively low-cost media. Additionally, there are many vector systems available for yeast expression and genetic manipulation, particularly for *S. cerevisiae*. A broad range of auxotrophic markers for selection of genetically engineered *S. cerevisiae* are also available [33].

Lastly, an oleaginous yeast *Yarrowia lipolytica* has been discovered to be able to utilize hydrophobic substrates (HS) like alkanes, fatty acids and oils for its growth. This yeast can solubilise HS by secreting lipases and uptake HS through protrusions on their cell surface. They also possess metabolic pathways to degrade alkanes and triglycerides [34]. More recently, hydrophobic/alkane-binding protein complexes have been identified in *Y. lipolytica* [35]. Studies to investigate genes responsible for alkane uptake, transport, metabolism and their regulation might prove to be useful for this study. For instance, with *Y. lipolytica* and *S. cerevisiae*-both being yeast species that-are more genetically close to each other, the possibility of introducing alkane-regulating genes from *Y. lipolytica* into *S. cerevisiae* could help to further improve alkane production.

2.3 Fatty acid (FA) as a precursor in the alkane biosynthesis pathway

The FAR acts on an acyl-CoA substrate, reducing it to aldehyde before finally being converted to alkane by AD. Acyl-CoA is formed by coenzyme A-activation of FAs in the cellular FA biosynthesis pathway. Thus, FA is an important precursor in alkane biosynthesis. In yeast, cellular FAs can be obtained from (i) exogenous source, (ii) endogenous lipid or protein turnover or (iii) *de novo* FA synthesis and elongation (Figure 5) [36].

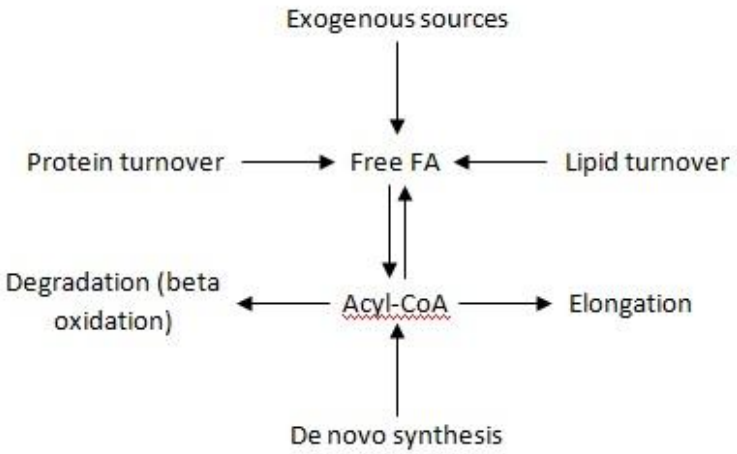


Figure 5 Schematic representation of FA metabolism in *S. cerevisiae*

In the initial step of FA synthesis, a precursor acetyl-CoA is carboxylated to malonyl-CoA. Carbon dioxide and ATP are required for this reaction. This malonyl-CoA serves as a two-carbon donor in a FA elongation cycle catalyzed FA synthase complexes and elongases which consist of several reductases and dehydratases (Figure 6).

In yeast, there are two types of fatty acid synthase (FAS): cytosolic FAS (type I) which consists of subunits Fas1 (β) and Fas2 (α) and mitochondrial FAS (type II). FA elongation

takes place in endoplasmic reticulum and follows similar pathways as FAS. The two most abundant FAs in yeast are C_{16} and C_{18} FAs, with C_{26} being the longest. Medium-chain FAs that are of great interest in this study-like C_{10}, C_{11} and C_{12}-are not naturally present and therefore, must be uptaken from external sources.

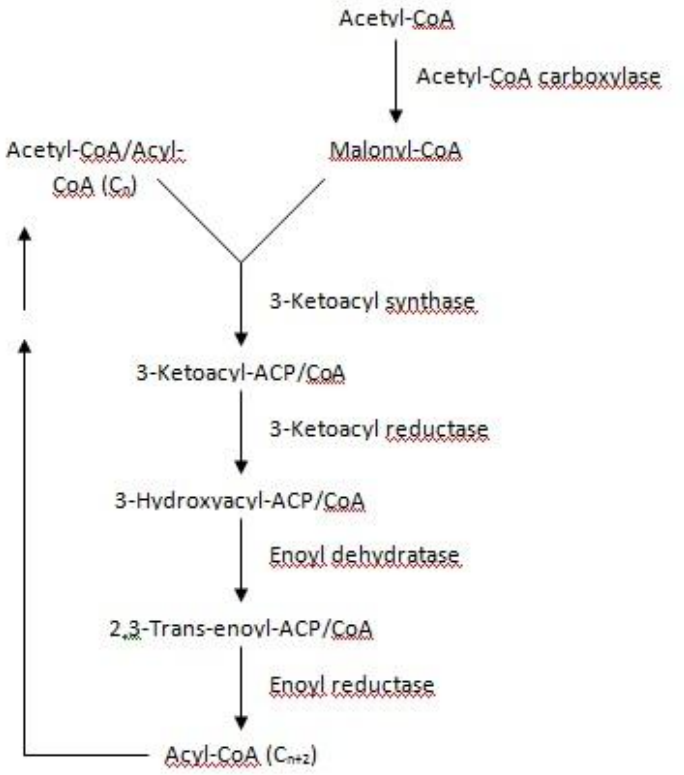


Figure 6 Schematic representation of FA synthesis and elongation in *S. cerevisiae*

The main degradation pathway of cellular FAs is the beta oxidation pathway, a reaction cycle in which two carbons are removed from acyl-CoA each turn (Figure 7). In yeast, beta oxidation occurs mainly in the peroxisomes and *POX* (*FOX*) genes have been identified to be responsible for this beta oxidation of FA. Three genes that are crucial in the beta oxidation cycle are *POX1* (acyl-CoA oxidase), *POX2* (enoyl-CoA hydratase) and *POX3* (3-ketoacyl-CoA thiolase) [37, 38]. Accumulation of FAs inside yeast cells can be made possible by deletion of

these genes, thus increasing the driving force for acyl-CoA reduction into aldehyde and eventually, the production of alkane.

2.4 Possible competing pathways involving aldehyde metabolism in S. cerevisiae

In *S. cerevisiae*, pathways that metabolize aldehyde mainly involve genes that are responsible for cell resistance towards toxic aldehydes such as formaldehyde, acetaldehyde and furfural. There are seven known alcohol dehydrogenases (Ald1p-7p) in *S. cerevisiae* that can reduce these aldehydes to alhcohol products. Short chain aldehydes like formaldehyde and acetaldehyde are reduced to methanol and ethanol by alcohol dehydrogenase 1 (Adh1p) which is encoded by the gene *YOL086C* [39]. Medium and longer chain aldehydes are reduced by Adh6p (*YMR318C*) [40] and Adh7p (*YCR105W*) [41]. These enzymes catalyze the reversible conversion of alcohols to aldehydes/ketones with the help of redox cofactor NAD^+ or $NADP^+$. Although Adh enzymes can act on both alcohol and aldehyde substrates, kinetic study revealed that Adh6p and Adh7p are NADPH-dependent aldehyde reductases, with their reductive efficiency being 50- to 100-fold higher than the corresponding oxidative reactions [40, 42].

Another way an aldehyde can be metabolized is its oxidation to fatty acid by an aldehyde dehydrogenase (ALD). In *S. cerevisiae*, acetaldehyde can be oxidized to acetate by two types of ALD: cytosolic and mitochondrial ALDs. Cytosolic ALDs are encoded by *ALD2*, *ALD3* and *ALD6*, and are $NADP^+$- and Mg^{2+}-dependent [43, 44]. Mitochondrial ALDs are encoded by *ALD4* and *ALD5* and are NAD^+/$NADP^+$- and K^+-dependent [45]. These genes are activated in response to stresses such as growth in ethanol-containing media. In summary, the competing pathways involving oxidation and reduction of aldehydes are shown in Figure 8.

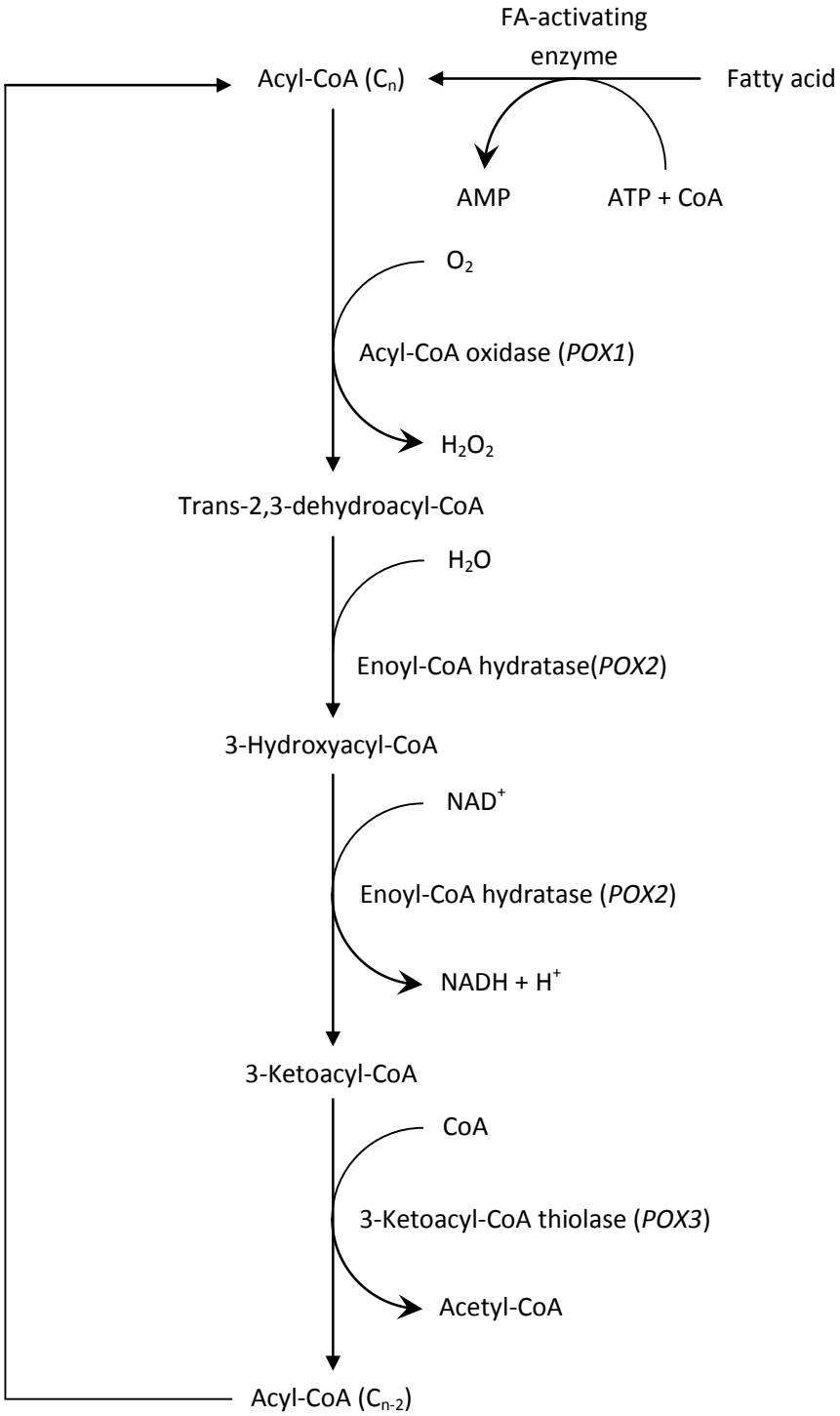


Figure 7 Schematic representation of FA beta oxidation pathway in *S. cerevisiae*

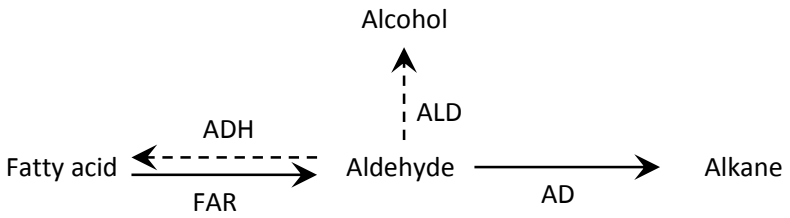


Figure 8 Schematic representation of aldehyde oxidation and reduction pathways in *S. cerevisiae*. Dashed arrows represent competing pathways in the alkane biosynthesis pathway.

3. Materials and Methods

3.1 Cloning of aldehyde decarbonylase (AD) genes into *E. coli* Top10

AD genes are cloned into *E. coli* Top10 for recombinant gene amplifications. Nucleotide sequence encoding *Anabaena variabilis* ATCC 29413 Ava_2533 (renamed as AD1 throughout this thesis, accession no. YP_323043) was synthesized and purchased from GeneArt (Life Technologies, Singapore) and cloned into the KpnI (3') and XhoI (5') sites of pYES2/CT, a *S. cerevisiae* expression vector. The vector contains yeast *GAL1* (galactose-inducible) promoter, C-terminal polyhistidine (6×His) tag for detection and purification of recombinant proteins, *URA3* auxotrophic marker for selection of yeast transformants and ampicillin resistance gene for selection of *E. coli* transformants. Nucleotide sequences encoding *Arthrospira platensis* str. Paraca AplaP_00840 (AD2, ZP_06380209) and *c* (AD3, YP_400610) were also synthesized and purchased from GeneArt and cloned into the KpnI (3') and XhoI (5') sites of the pYES2/CT vectors. All genes were codon optimized for expression in *S. cerevisiae*. The three constructs were individually transformed into *E. coli* Top10 (Invitrogen, Singapore) at 42°C and the cells were grown at 37°C in LB solid medium containing 10 g/l tryptone, 5 g/l yeast extract, 10 g/l NaCl, 15 g/l agar and 100 µg/ml ampicillin for selection of transformants. The plasmid DNA encoding AD1 from the recombinant colonies was PCR-amplified using primers pyes2ct-f and pyes2ct-r (see Appendix A) and the size of amplicons were then compared with a DNA ladder in a 1% agarose DNA gel electrophoresis to check for correct inserts' sizes. Plasmid DNAs encoding AD2 and AD3 were also checked using identical methods.

The plant decarbonylase genes *Cer1* from *Arabidopsis thaliana* and *Hvcer1* from *Hordeum vulgare* were also cloned in the yeast expression vector pYES2/CT. Insect cytochrome P450

CYP4G2 (G2) decarbonylase gene from *Musca domestica* was cloned into the yeast pESC-URA vector. All three genes were transformed into *E. coli* using the method described above.

3.2 Heterologous expression of AD genes in *S. cerevisiae* BY4741

Recombinant cells were grown at 37°C, 200 rpm in LB medium containing 10 g/l tryptone, 5 g/l yeast extract, 10 g/l NaCl and 100 µg/ml ampicillin. Plasmid DNAs containing the six genes were purified using QIAprep Spin Miniprep Kit (QIAGEN, Singapore) according to the manufacturer's protocol and transformed into *S. cerevisiae* BY4741 at 42°C. Yeast transformants were grown at 30°C in YNB solid medium containing 6.7 g/l Yeast Nitrogen Base without amino acids (BD Difco, Singapore) and 20 g/l agar, supplemented with 2% dextrose as carbon source and 0.77 g/l URA DO supplement (Clontech, Singapore) for yeast auxotrophic marker selection. Plasmid DNAs encoding AD1, AD2, AD3, Cer1p and Hvcer1p from recombinant colonies were PCR-amplified using primers pyes2ct-f and pyes2ct-r and the size of amplicons were then compared with a DNA ladder in a 1% agarose DNA gel electrophoresis to check for correct sizes. G2p gene size was checked using a pair of gene-specific primers.

Resulting yeast colonies were grown overnight at 30°C, 250 rpm in YNB medium containing 0.67 g/l Yeast Nitrogen Base w/o amino acids, supplemented with 2% dextrose and 0.77 g/l URA DO supplement. Cells from overnight cultures were induced in YNB medium supplemented with 2% galactose and 0.77 g/l URA DO supplement at OD_{600} of 0.4. The cultures were further grown for another 24 h at 30°C, 250 rpm, then harvested by centrifugation at 5400×g for 5 min and washed with H_2O.

3.3 mRNA identification by reverse transcription PCR (RT-PCR)

To extract the total RNA, yeast cells were first mechanically lysed using glass beads in a FastPrep lyser (MP Biomedicals, Singapore). Total RNA extraction was then carried out using RNeasy Mini Kit (QIAGEN, Singapore) according to the manufacturer's protocol. Purified RNA was then incubated with RevertAid H Minus Reverse Transcriptase (Fermentas, Singapore) to create cDNA from mRNA, and the resulting cDNA was used as a template for subsequent PCR. PCR products were then run on a DNA agarose gel electrophoresis to check for correct sizes.

3.4 Purification of AD1, AD2, AD3 and *in vitro* aldehyde decarbonylase assay

Pelleted cells from 200 ml of induction culture were resuspended in 10 ml (approx. 100 mg wet wt. of cells/ml) of lysis buffer (Tris) of the following composition: 50 mM Tris-HCl, 150 mM NaCl, 10 mM EDTA, 0.5% Tween 20, serine protease inhibitor (PMSF), pH 7.5-8.0. Aliquots of the resuspended mixture were then added to 1.5-ml microtubes containing equal volume of glass beads and lysed at 4.0 m/s for 3×40s (1 min rest in between cycles) in a FastPrep. The insoluble material was pelleted by centrifugation at 4000×g for 15 min. The supernatant was collected and buffer-exchanged using a Sephadex G-25 column (GE Healthcare, Singapore) into binding buffer with the following composition: 50 mM NaH_2PO_4, 300 mM NaCl, 5 mM imidazole, pH 8.0. Purifications of the His-tagged recombinant proteins were achieved by Ni-affinity chromatography using a GE AKTA Explorer fast protein liquid chromatography (FPLC) system (GE Healthcare, Singapore) equipped with a HisTrap HP column (GE Healthcare). Bound proteins were washed with binding buffer and eluted in a step-gradient at 1 ml/min with an elution buffer comprising: 50 mM NaH_2PO_4, 300 mM NaCl,

250 mM imidazole, pH 8.0. After elution, the purified proteins were buffer-exchanged into a 0.1 M sodium phosphate assay buffer, pH 7.2. Protein expression and purified fraction purity were determined by SDS-PAGE and protein immunoblot (western) analyses while protein concentration (Appendix B) was determined using Coomassie (Bradford) Protein Assay Kit (Thermo Scientific, Singapore) according to the manufacturer's protocol. Coomassie Blue staining was used to visualize protein bands in SDS-PAGE whilst 3,3',5,5'-tetramethylbenzidine (TMB) liquid substrate system for membranes (Sigma-Aldrich, Singapore) was used for visualization of protein bands on western blot PVDF membrane (BioRad, USA), together with anti-6×His monoclonal antibody-horseradish peroxidase (HRP) conjugate (Clontech, Singapore).

Enzymatic assays were then performed in a 500 µl reaction volume. A typical enzymatic reaction consisted of up to 20 µM AD, 40 µM ferrous ammonium sulphate in the form of $(NH_4)_2Fe(SO_4)_2 \cdot 6H_2O$ (Sigma, Singapore), 75 µM phenazine methosulfate (PMS), 1 mM nicotinamide adenine dinucleotide (NADH), 20 µM bovine serum albumin (BSA) and various concentrations of dodecanal in the assay buffer. Dodecanal 500× stock solutions were prepared in DMSO. The reactions were performed at 37°C, 650 rpm for approx. 30 min and then quenched by addition of 300 µl of ethyl acetate and vortexed vigorously for 15 min to extract reaction product and unreacted substrate. 1-Octadecene was used as an internal standard at a concentration of 5 mg/l. The organic layer was separated by centrifugation at 13000 rpm for 5 min. A 1 µl sample of ethyl acetate layer was then injected into the GC-MS column for analysis of hydrocarbons.

3.5 Quantification and analysis of hydrocarbon by GC-MS

GC-MS analysis of hydrocarbons was performed using an Agilent 7890A GC system coupled with 5975C inert MSD. Agilent HP-5ms column (30 m × 250 µm × 0.25 µm) was used with helium carrier gas at a flow rate of 1.1 ml/min. Inlet temperature was 250°C and 1 µl of the sample was injected in splitless mode. Oven temperature was held at 50°C for 0.5 min, increased to 280°C at 20°C/min and maintained at 280°C for 3 min, with a total run time of 15 min. The MS source temperature was 230°C and the solvent delay was 3 min. Undecane was eluted at 4.95 min while 1-octadecene was eluted at 9.5 min.

3.6 *In vivo* AD assay

Resulting yeast colonies were grown overnight at 30°C, 250 rpm in YNB medium, supplemented with 2% dextrose and 0.77 g/l URA DO supplement. Cells from overnight cultures were induced in YNB medium supplemented with 2% galactose, 50 µM $(NH_4)_2Fe(SO_4)_2 \cdot 6H_2O$ and 0.77 g/l URA DO supplement at OD_{600} of 0.4 for 24 h, after which various concentration of aldehydes were added from stocks. Dodecanal and tetradecanal 20 mM stock solutions were prepared in 10% Triton X-100, while 40 mM hexadecanal and octadecanal stocks were prepared in isopropanol. The cultures were further grown for another 24-48 h at 30°C, 250 rpm, then harvested by centrifugation at 5400×g for 5 min and washed with H_2O. In this step, the growth medium was collected for hydrocarbon analysis. To extract hydrocarbons, 300 µl of ethyl acetate was then added into either cell pellet or growth medium and the mixture was vortexed vigorously for 1 h. Following centrifugation at 13000×g, 5 min to separate the organic and aqueous layers, the sample was injected into the GC-MS for identification of hydrocarbons.

4. Initial Investigation of Decarbonylase Activity: *In Vitro* Assay

In order to investigate the activity of several types of decarbonylases, ADs from cyanobacteria and plants were selected and cloned into *S. cerevisiae*. Several of these enzymes were then expressed recombinantly and purified for *in vitro* activity assay.

4.1 Selection of ADs from *Synechococcus elongatus*, *Anabaena variabilis*, *Arthrospira platensis*, *Arabidopsis thaliaha* and *Hordeum vulgare*.

Since PCC7942_1593 (AD3) has been reported to have decarbonylase activity and confer alkane biosynthesis in *E. coli*, AD3 was chosen as a benchmark candidate in this study. cADs from *A. variabilis* (AD1) and *A. platensis* str. Paraca (AD2) were selected because the genomes of both cyanobacteria have been fully sequenced and both AD1 and AD2 have a high amoni acid sequence identity, i.e. 74% and 78%, respectively, with AD3. Furthermore, *A. variabilis* has been reported to synthesize hydrocarbons. When photoautotrophically grown and its culture extract analyzed for hydrocarbons, *n*-heptadecane, 7-methyl-heptadecane and 8-methyl-heptadecane were detected [1, 46]. These alkanes accounts for approximately 0.09% of the cellular dry weight.

Two putative ADs from plants have also been selected and cloned into *S. cerevisiae*, in an attempt to achieve heterologous expression and protein extraction for enzyme characterization. The two AD genes were *Cer1* and *Hvcer1* from *A. thaliana* [3] and *H. Vulgare* [12], respectively. To verify that the two yeast clones carry the correct *Cer1* and *Hvcer1* genes, two sets of primers were designed for PCR: one set of vector-specific primers and one set of gene-specific primers (Figure 9, Appendix A). The colony PCR products were

then examined for the presence of *Cer1* and *Hvcer1* genes in the two clones by DNA electrophoresis (Figure 10). Although these genes have been successfully cloned into *S. cerevisiae*, protein expression and purification proved difficult. SDS-PAGE analysis, western blot and RT-PCR results indicated very low or no expression. While RNA extraction was successful, no DNA bands were detected by RT-PCR (Figure 11), indicating that the mRNAs encoding Cer1p and Hvcer1p were highly unstable or present at an exceptionally low level. This result is consistent with analyses of Cer1p and Hvcer1p expression, where no protein bands were detected by SDS-PAGE and western blot. A possible explanation is because these genes are large membrane proteins found in leaves, stalks and other external surfaces in plants, making their soluble expression in *S. cerevisiae* a challenging task. A Transmembrane Predictor (TMHMM) software (see Section 4.3 below) predicted that these proteins have 4 transmembrane helices that span the cell membrane.

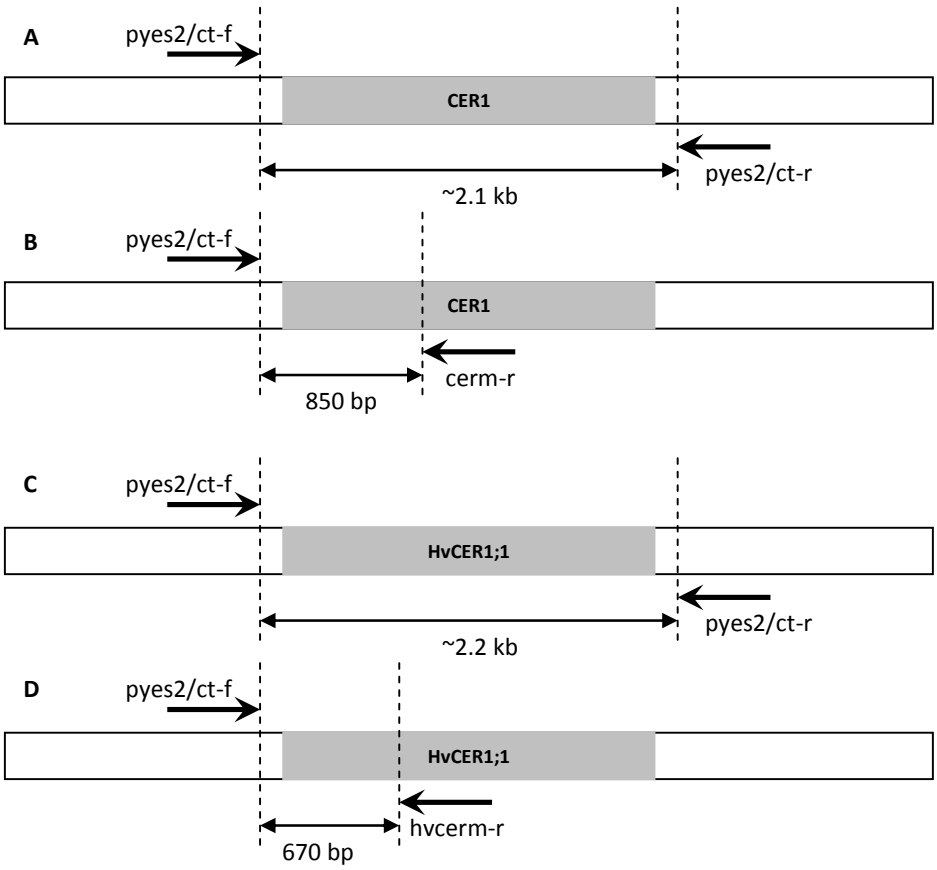


Figure 9 Schematic diagram of PCR amplification of *Cer1* and *Hvcer1* genes. (A) PCR amplification of *Cer1* using pYES2/CT-F and pYES2/CT-R (vector-specific) primers where the product is the whole CER1 gene, approx. 2.1 kb. (B) Replacing pYES2/CT-R with gene-specific primer Cerm-R results in shorter product, 850 bp. (C) Similarly, *Hvcer1* amplification using vector-specific primers results in a PCR product of approx. 2.2 kb. (D) Replacing pYES2/CT-R with Hvcerm-R results in a 670 bp fragment.

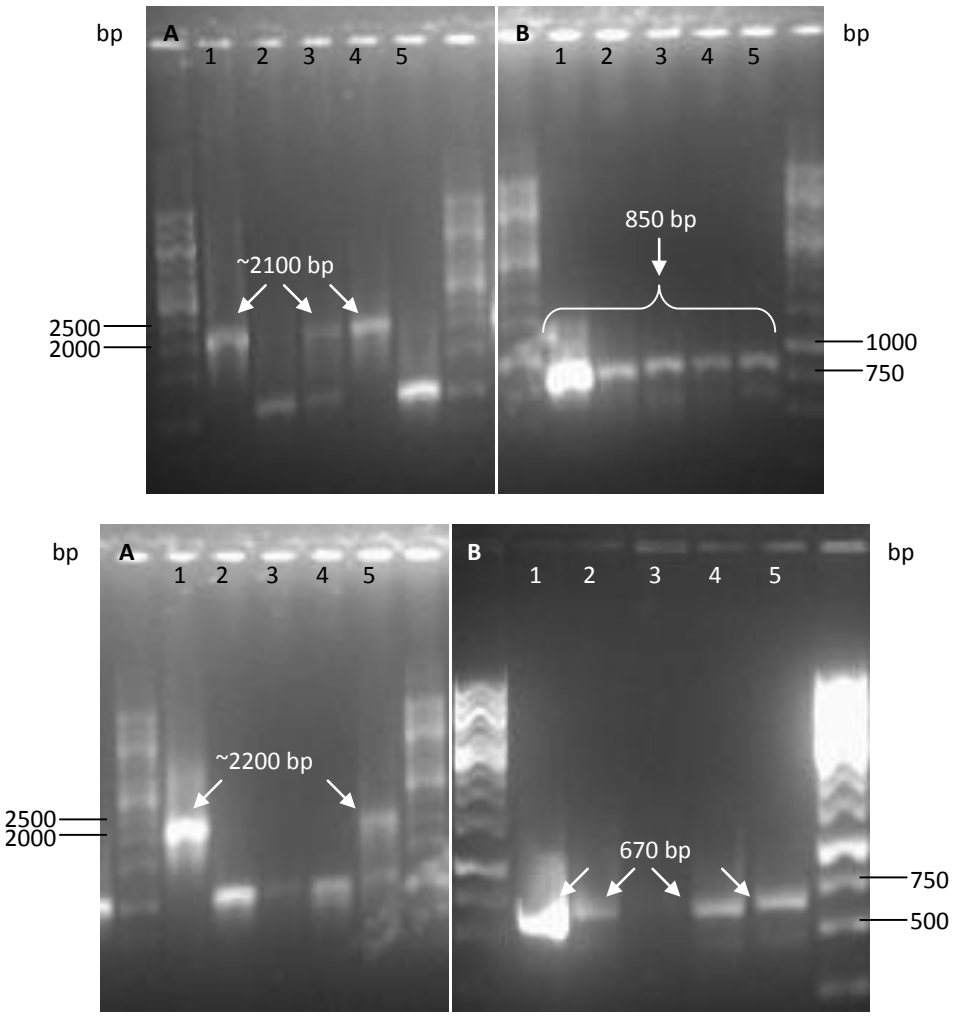


Figure 10 Yeast colony PCR products of *Cer1* and *Hvcer1* genes. (A) *Cer1* using vector-specific primers (B) *Cer1* using gene-specific primers. (C) *Hvcer1* using vector-specific primers. (D) *Hvcer1* using gene-specific primers. Correct bands are indicated by white arrows. Lane 1 is a positive control (plasmid with CER1 gene). Lanes 2-5 are PCR products from different colonies.

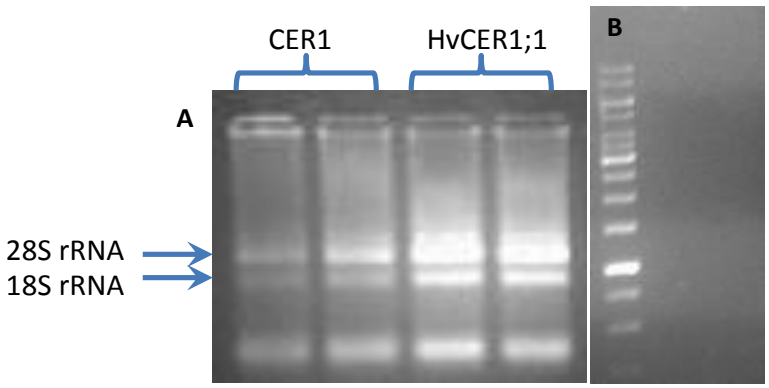


Figure 11 (A) Yeast total RNA, extracted using QIAGEN RNeasy Mini Kit. (B) RT-PCR result of Cer1 and Hvcer1 mRNAs. No bands were visible.

4.2 Cloning and heterologous expression of AD1, AD2 and AD3

The three genes, AD1, AD2 and AD3, were successfully cloned into the expression vector pYES2/CT and transformed into *E. coli* Top10. This result was confirmed by bacterial colony PCR and DNA agarose gel electrophoresis (Figure 12). The size of PCR amplicons are 985 base pairs (bp), 985 bp and 988 bp for AD1, AD2 and AD3, respectively.

Similarly, successful transformations of the genes into *S. cerevisiae* were confirmed by yeast colony PCR and DNA gel electrophoresis (Figures 13A and B). We also identified the existence of AD3 mRNA in the yeast by RT-PCR, signifying that transcript of the recombinant genes PCC7942_1593 is present (Figure 13C). Unlike bacterial colony PCR, DNA bands from yeast colony PCR have considerably less intensity. This may be due to the fact that yeast cells are much more difficult to break. Plasmid DNAs are not completely released from yeast cells due to the incomplete lysis, causing them inaccessible for primer annealing and extension during PCR. To address this constraint, up to 2% Triton X-100, a non-ionic surfactant, was added into polymerase buffer to improve cell lysis and DNA denaturation. No products were obtained in yeast colony PCR without the use of Triton X-100, as identified using DNA agarose gel electrophoresis.

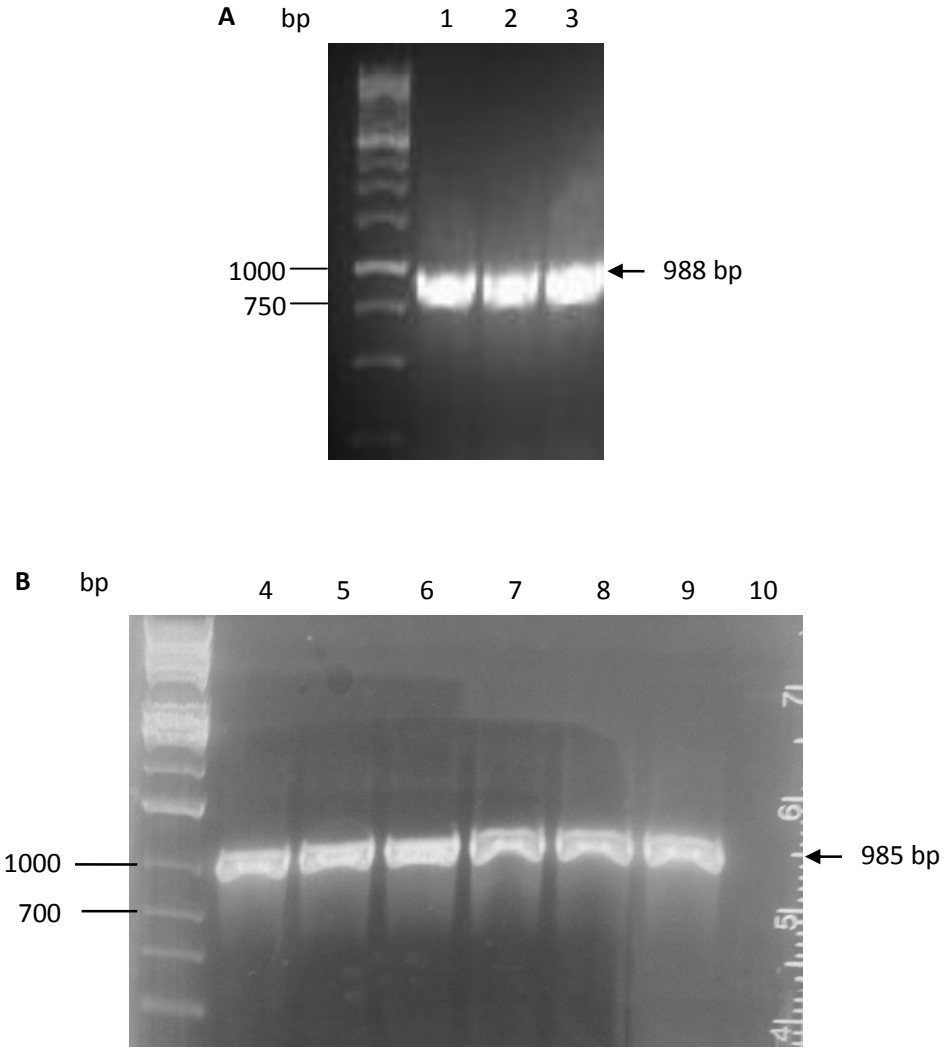


Figure 12 DNA gel electrophoresis of bacterial colony PCR amplicons (A) AD3 (B) AD1 and AD2. Lanes 1,2,3: colonies of AD3; lanes 4,5,6: colonies of AD1; lanes 7,8,9: colonies of AD2; and lane 10: negative control. Bands at 985 bp indicate the correct gene size.

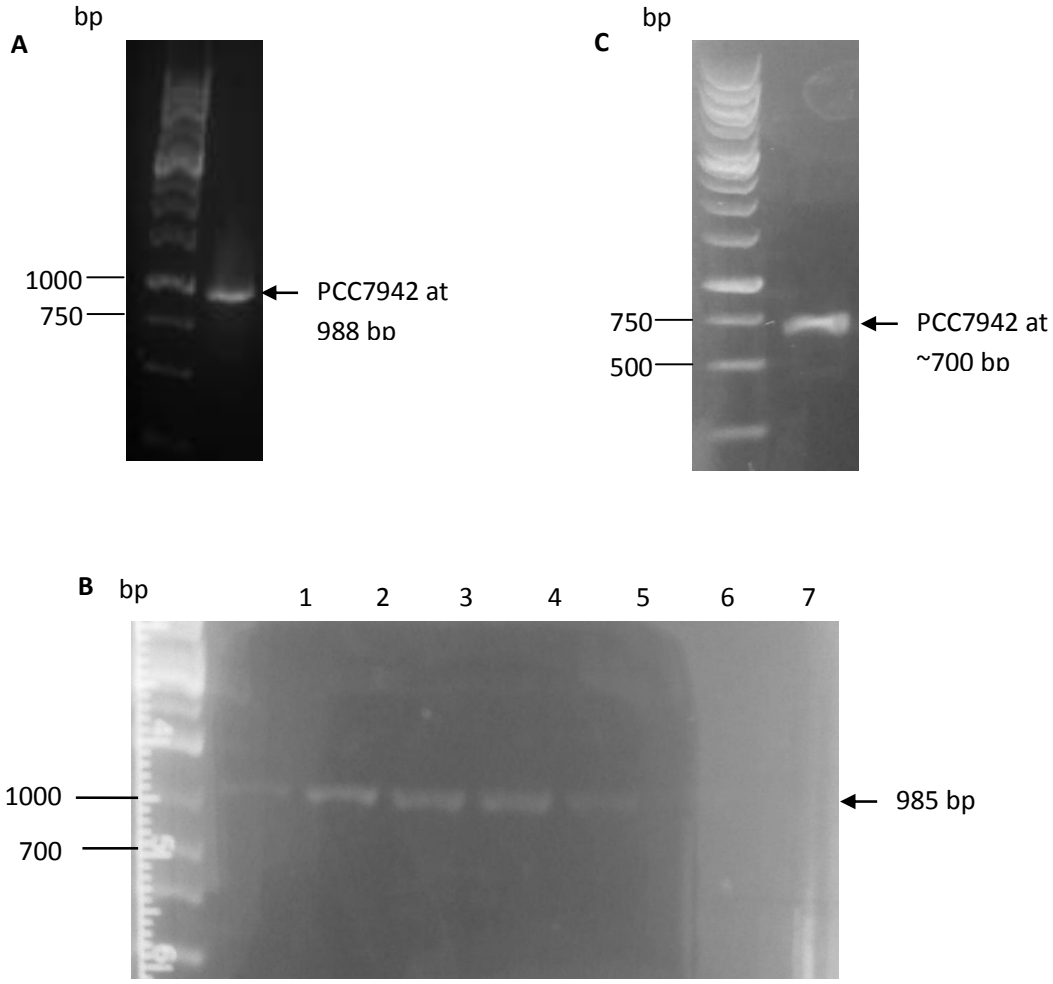
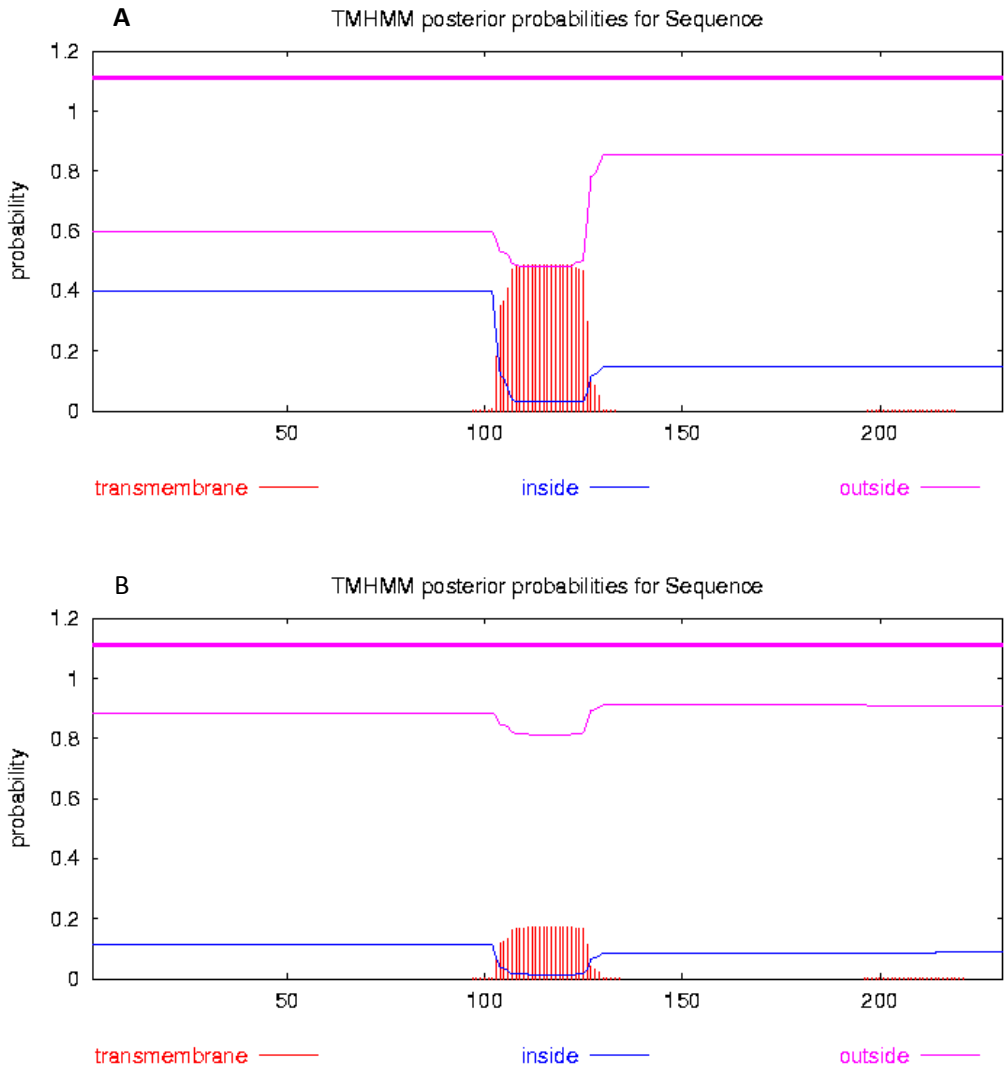


Figure 13 (A) DNA gel electrophoresis of yeast colony PCR products for AD3, (B) for AD1 (lanes 1,2,3), AD2 (lanes 4,5,6) and a negative (lane 7) and (C) for RT-PCR of AD3.

4.3 Purification of recombinant cADs

Recombinant cADs were purified for *in vitro* enzyme characterization. In order to design the purification strategy, it is important to first establish the location of the cADs in the cell. This was done by employing the Transmembrane Helices Predictor (TMHMM) software (TMHMM Server v. 2.0, www.cbs.dtu.dk/services/TMHMM-2.0), which estimates the probability that a segment of a protein is located inside or outside the plasma membrane, or is transmembrane. Our initial prediction of the three enzymes using the TMHMM software

showed that all three cADs do not have transmembrane helices. All three cADs show similar probability profiles, where they were predicted to be located outside the membrane with amino acids (AA) in the middle of the sequence (between position 100-150) having increased possibilities to be transmembrane helices (Figure 14). This result suggests that they might be membrane-associated proteins (peripheral membrane proteins). AD1 has the highest possibility as AA residues between position 100 and 125 have a probability of 0.50 probabilities to be located in the membrane and 11 out of its 231 AA residues are predicted to be transmembrane helices. AD2 has 4 AAs while AD3 has 6 AAs in the transmembrane helices.



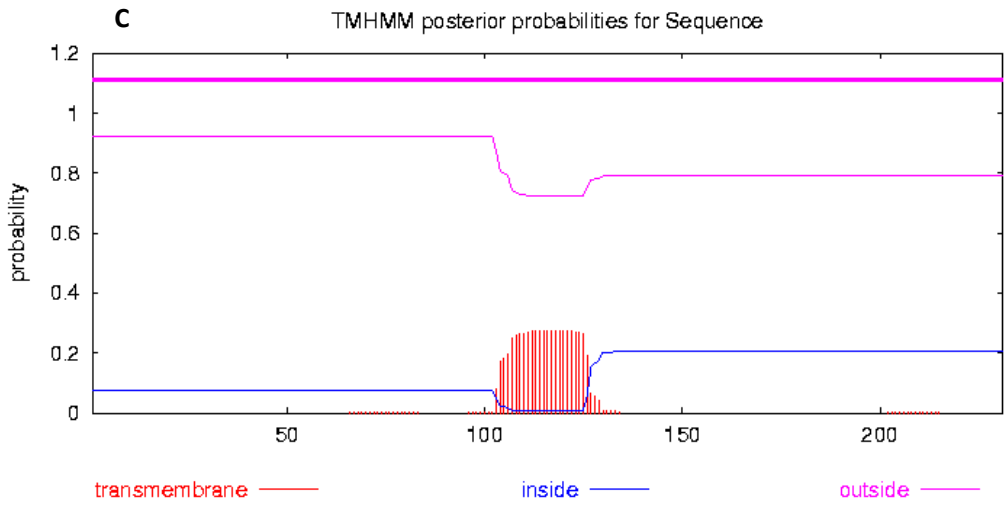


Figure 14 Probability plots of transmembrane helices in cADs as predicted by TMHMM. (A) AD1, (B) AD2 and (C) AD3.

To achieve the best condition that will yield optimal purification of these cADs, several lysis buffers were used for extraction of AD3 as a representative of the other cAD candidates. The buffers studied were (i) Buffer A (a denaturing buffer composed of 4% SDS, 10% glycerol, 5% 2-mercaptoethanol and 62.5 mM Tris-HCl, pH 6.8), (ii) Buffer B and (iii) Buffer C (non-denaturing buffers composed of 50 mM Tris-HCl, 150 mM NaCl, 10 mM EDTA and either 2% Triton X-100 for Buffer B or 0.05% Tween 20 for Buffer C, pH 7.5-8.0).

Buffer A was used for yeast total protein extraction to detect the presence of AD3 in the cell lysate. AD3 protein could be easily extracted and purified using Buffer A, as determined by SDS-PAGE analysis (Figure 16). However, the resulting protein was in a denatured form and thus, not suitable for activity assay. To obtain purified proteins in their native form, Buffer A was substituted with either Buffer B or C. The target protein was detected in the soluble fraction of cell lysate and in the purified fraction (Figure 15) when Buffer B or C was used, as determined from western blot analysis (Figure 17C). More AD3 was recovered using Buffer C extraction compared to Buffer B, indicating that a lower concentration of Tween 20 (0.5% versus 2% of Triton X-100) was sufficient to extract larger amounts of AD3. Furthermore, the

higher concentration of Triton X-100 and the fact that it is difficult to be removed, even after buffer exchange [47], might cause more interference with the binding to Ni-column for IMAC purification, hence impairing the purification. The SDS-PAGE analysis, however, shows a significant decrease in the amount of extracted protein using Buffer B or C compared to that extracted with Buffer A. While Buffer B or C can extract an amount of protein high enough to be detected using western blot, it is not sufficient for SDS-PAGE detection (Figure 18C).

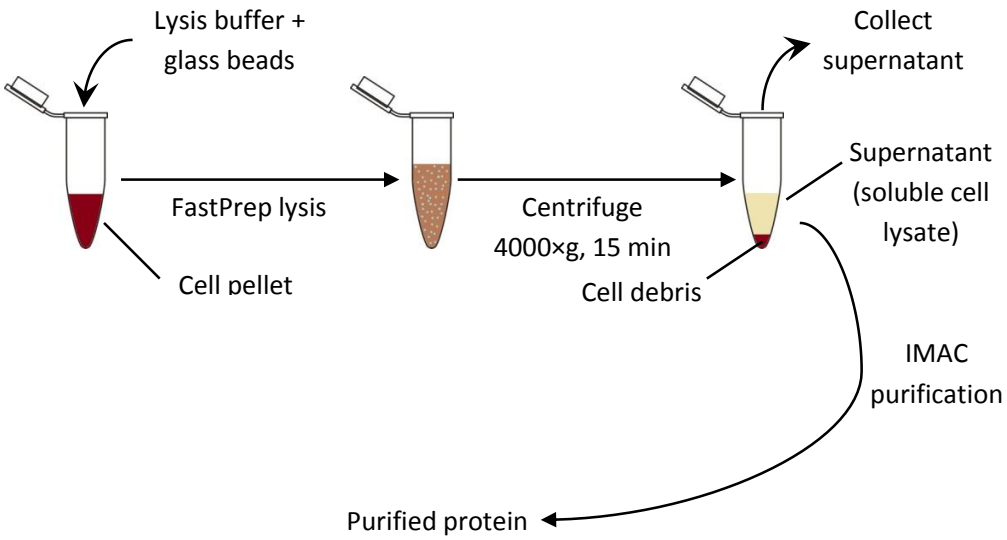


Figure 15 Schematic diagram of cAD purification process

Similar results were obtained when Buffers B and C were used for AD1 and AD2 extractions. For both cADs, target protein bands were thicker when Tween 20 was used and again, the use of Triton X-100 resulted in co-extraction of non-target proteins in purified fractions, indicative of Ni-column binding interference (Figures 17A, B and 18A, B). Interestingly, the use of Buffers B and C resulted in higher protein extraction yield for AD1 and AD2 than AD3, as determined by SDS-PAGE analyses. The AD1 and AD2 protein bands when extracted using Buffer B or C also showed higher intensity than AD3 band when extracted using Buffer A.

Comparison of western blot results of AD2 and AD3 also shows similar result in both soluble cell extract and purified fractions. This could be due to the lower expression level of AD3 compared to AD1 and AD2.

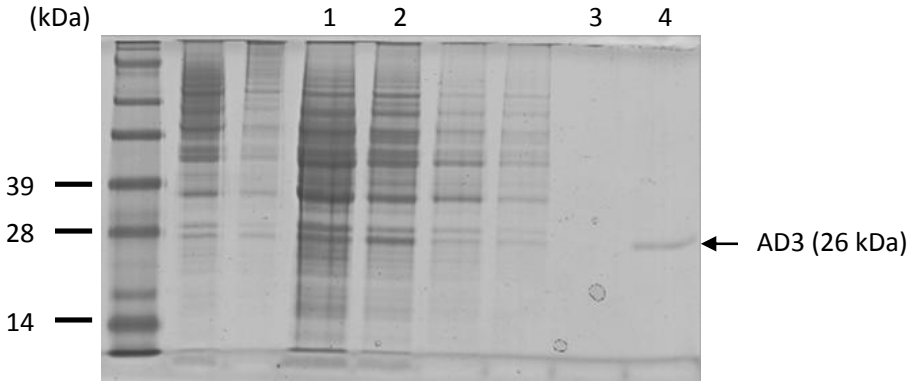


Figure 16 SDS-PAGE analysis of AD3 extraction in Buffer A. Lane 1: control, soluble cell lysate; lane 2: soluble cell lysate; lane 3: control, purified fraction; lane 4: purified fraction.

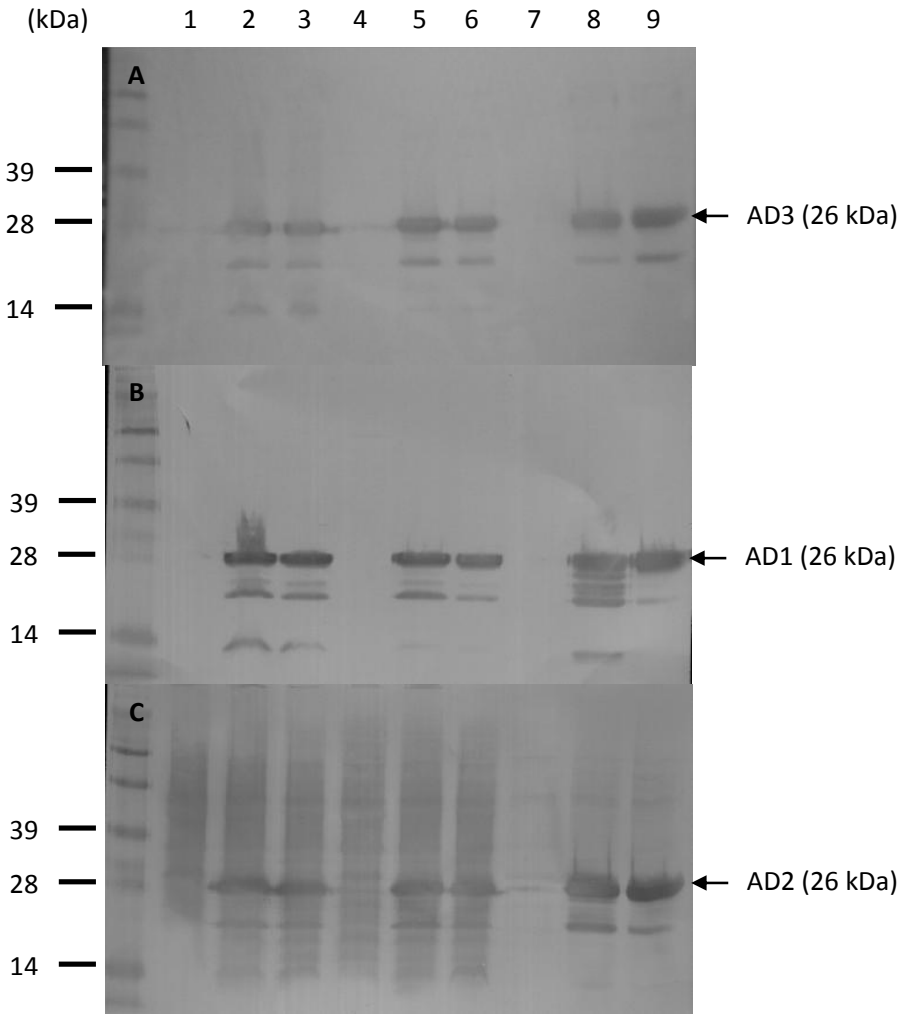


Figure 17 Western blot analysis of (A) AD3, (B) AD1 and (C) AD2. Lanes 1-3: crude lysates; lanes 4-6: soluble cell lysates; lanes 7-9: purified fractions. Lanes 1, 4, 7: control; lanes 2, 5, 8: Buffer B extraction; lanes 3, 6, 9: Buffer C extraction.

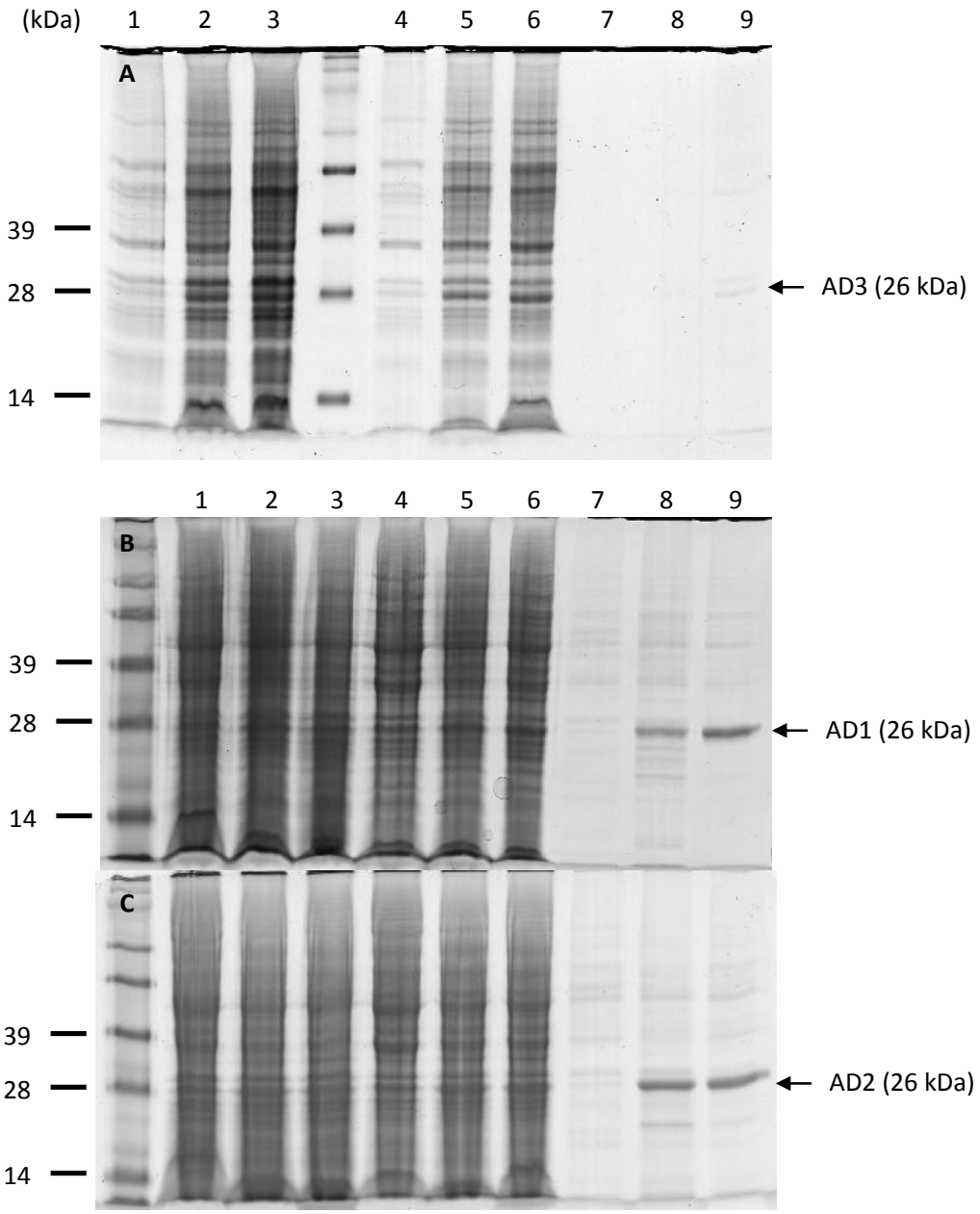


Figure 18 SDS-PAGE analysis of (A) AD3, (B) AD1 and (C) AD2. Lanes 1-3: crude lysates; lanes 4-6: soluble cell lysates; lanes 7-9: purified fractions. Lanes 1, 4, 7: control; lanes 2, 5, 8: Buffer B extraction; lanes 3, 6, 9: Buffer C extraction.

We conclude that Buffer C is the most optimal buffer for extraction of cADs in their native forms as Tween 20, a polyoxyethylene surfactant, has larger polar headgroups and more water-soluble [48]. Hence, Tween 20 is milder than Triton X-100. The purified AD1 and AD2

had purities ≥ 95% when extracted using Buffer C, as determined from SDS-PAGE analysis and Gel-Pro Analyzer version 3. The purities when Buffer B was used were 86% and 84% for AD1 and AD2, respectively, while AD3 purity when extracted with Buffer A was also ≥ 95%. As AD3 expression level was low and could not be extracted in its native form using Buffer B or C, AD3 was not included in the enzyme kinetic studies.

4.4 *In vitro* enzymatic assay

The activity of cADs was characterized using an *in vitro* cAD assay as described in Section 3.4, except that in one study the PMS/NADH auxiliary reducing system was substituted with ferredoxin/ferredoxin reductase/NADPH to examine the effectiveness of the latter system. Enzyme kinetics studies were performed using PMS/NADH. The enzyme assay results show that, though very low, AD1, AD2 and AD3 showed bioactivity, which confirms that their heterologous expression in *S. cerevisiae* did not compromise enzyme activity. Importantly, the presence of BSA was critical in improving cAD enzyme activity, where increased production of undecane by 46%, 90% and 87% for AD1, AD2 and AD3, respectively, was observed when BSA was included in the assay mixture (Figure 19). One possible explanation is that since BSA has been reported to solubilise hydrophobic lipids to a certain extent [25, 26], BSA could play a role in delivering the water-insoluble aldehyde substrate to the enzyme, making the substrate more accessible for conversion.

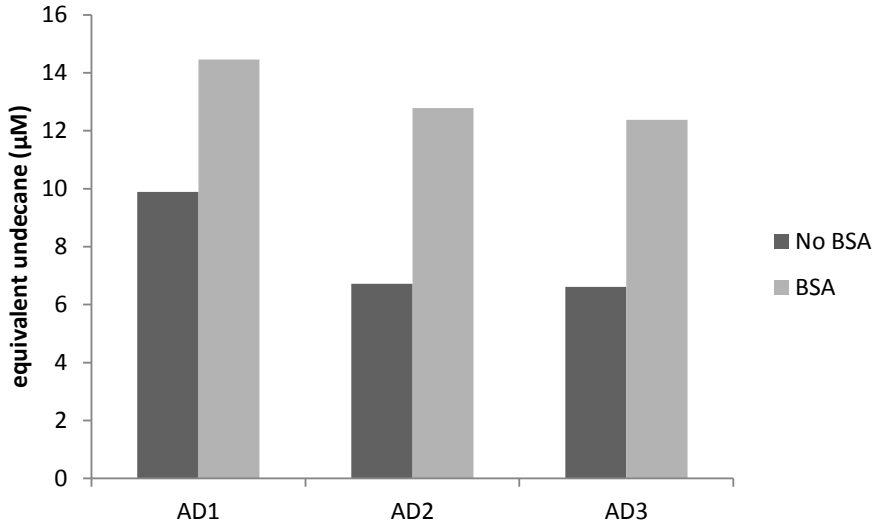


Figure 19 Undecane yield from *in-vitro* enzyme assay in ferredoxin/ferredoxin reductase/NADPH reducing system, with and without BSA. Undecane was not detected in a control without any cADs.

In vitro decarbonylase assays using the PMS/NADH reducing system also showed that the inclusion of BSA into the reaction increases the undecane yield by 143%, 61% and 50% for AD1, AD2 and AD3, respectively (Figure 20). In addition, the activity characterization results obtained also showed that without external reducing system or stoichiometric amount of Fe(II) (in the form of $(NH_4)_2Fe(SO_4)_2{\cdot}6H_2O$), the *in vitro* reaction will not take place.

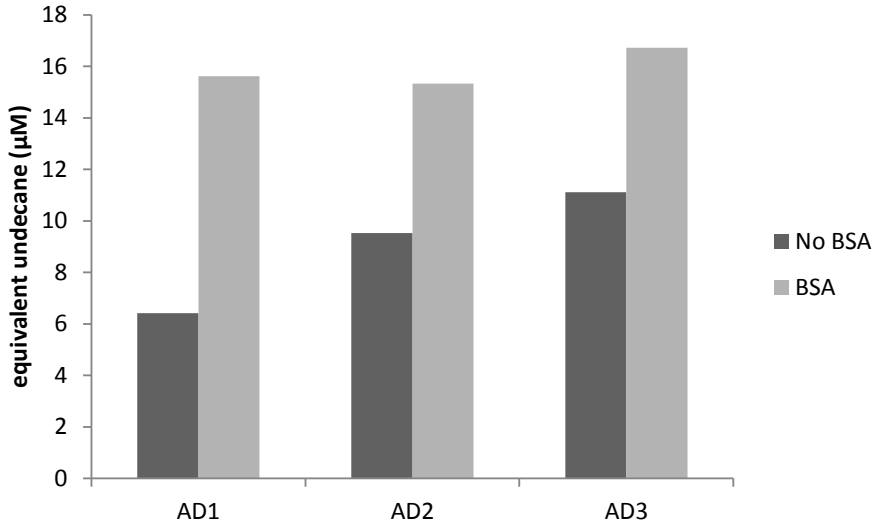


Figure 20 Undecane yield from *in-vitro* AD1 and AD2 assay in PMS/NADH reducing system, with and without BSA.

To quantitate the enzymatic behaviour of AD1 and AD2, the kinetic parameters for AD1 and AD2 under aerobic condition were determined using the Michaelis-Menten equation:

$$v = \frac{v_{max} \cdot [S]}{K_M + [S]};$$

where v is the rate of reaction, [S] is the substrate concentration and K_M is a constant.

The reaction rate (v) can be expressed as: $v = -\frac{d[S]}{dt}$. Substituting v with Michaelis-Menten equation and integrating from t = 0, [S] = $[S]_0$ to t, [S], the following can be obtained:

$$\int_{[S]_0}^{[S]} \frac{K_M + [S]}{[S]} d[S] = -v_{max} \int_0^t dt$$

$$[S(t)] = [S]_0 e^{-\frac{v_{max}}{K_M} t} \cdot e^{-\frac{([S]_0 - [S])}{K_M}}.$$

When only initial velocity is considered, [S] can then be calculated as: $[S(t)] = [S]_0 e^{-\frac{v_{max}}{K_M}t}$. Therefore, reaction rate $v = -\frac{d[S]}{dt} = \frac{[S]_0 \cdot v_{max}}{K_M} e^{-\frac{v_{max}}{K_M}\cdot t} = \frac{v_{max}}{K_M}[S]$. The rate constant k_{cat} is defined as:

$v = k_{cat}[ES]$ or $v_{max} = k_{cat}[E]_0$ when all enzyme is bound to substrate $[ES] = [E]_0$.

Hence, the reaction rate can be expressed as $v = \frac{k_{cat}}{K_M}[E]_0[S]$

Under the reaction conditions employed, as detailed in Section 3.4, these MM parameters can be determined by fitting experimental data into the MM equation (Figure 21), to obtain K_M = 3.5±2.9 μM and k_{cat} = 1.4±0.1/h for AD1 using dodecanal as the substrate. The K_M and k_{cat} values for AD2 are 5.0±2.2 μM and 0.9±0.1/h, respectively. Both AD1 and AD2 have extremely low activity, only approximately 1 turnover per hour. These findings are consistent with previous results [27, 29]. When decanal was used as a substrate for AD2, nonane was not detected in the product. Nonane, a C9 alkane, is also a medium chain alkane that is used for fuels.

One major challenge in this kinetics study is the insolubility of dodecanal substrate in aqueous solution. This causes complex reaction kinetics with substrate transfer to the enzyme as the rate limiting step. Hence, the parameters determined above might underestimate the real kinetics. The addition lipid-binding proteins, such as BSA, is one possible way to improve substrate transfer to the enzyme. Alternatively, the use of solvents that dissolve both polar and non-polar compounds or solvents that are miscible in both organic solvents and water could help improve substrate solubility.

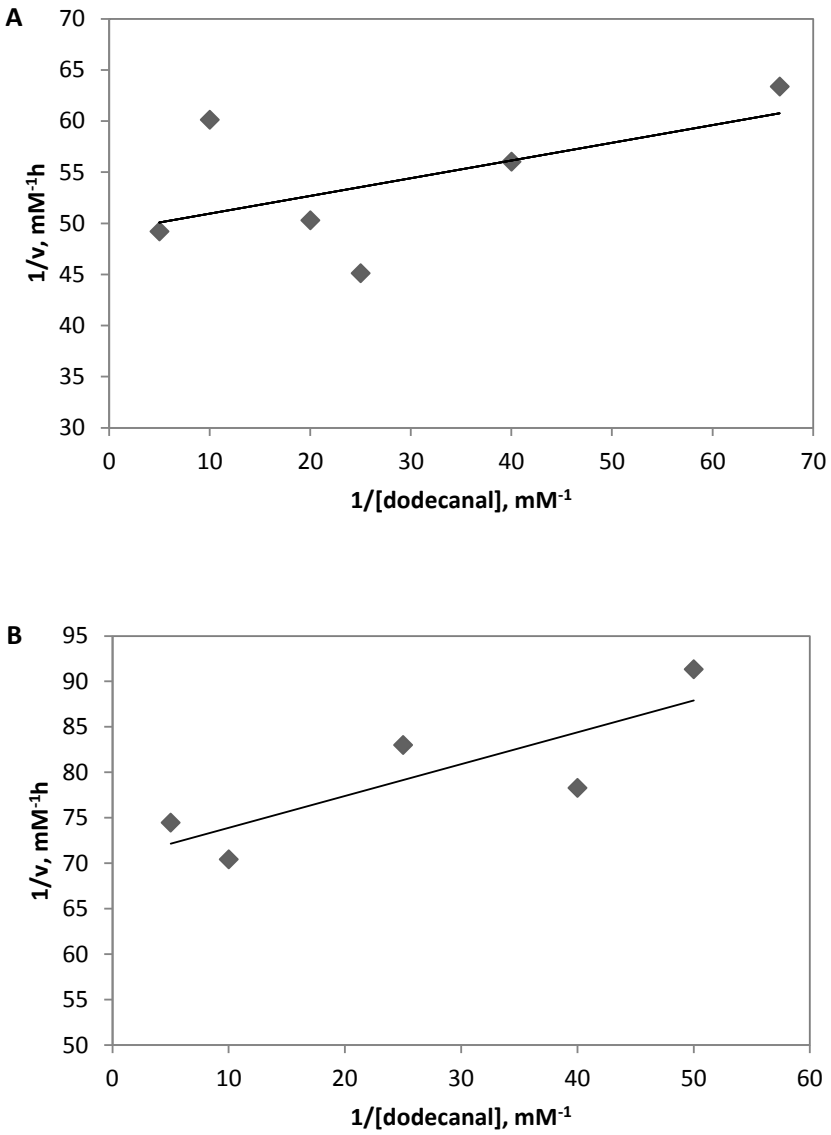


Figure 21 Lineweaver-Burk plots of Michaelis-Menten equation for (A) AD1, with K_M = 3.5±2.9 µM and k_{cat} = 1.4±0.1/h; and (B) AD2, with K_M = 5.0±2.2 µM and k_{cat} = 0.9±0.1/h. Reactions contained PMS/NADH (1 mM), BSA and dodecanal substrate. v is the rate of undecane formation ([undecane]/h). [] indicates concentration.

We next compared cAD activity when NADPH or NADH was used as a reductant. NADH plays a key role in catabolism reactions *in vivo* to generate ATP, where the cellular ratios of NAD^+/NADH are kept high to provide ample supply of NAD^+ as an oxidizing agent. NADPH, on the other hand, is mainly involved in anabolism reactions as an electron donor and the ratio of $NADP^+$/NADPH is relatively lower than that of NAD^+/NADH [49]. In *S. cerevisiae*, NADPH,

generated mainly from the pentose phosphate pathway [50], is readily available in the cell to act as a reducing agent. Therefore, the ability of NADPH to sustain cAD activity over NADH could be important for *in vivo* decarbonylation.

As NADPH is a more abundant reducing cofactor in *S. cerevisiae* cells compared to NADH [49], we examined the activity of cAD when incubated in the PMS/NADPH reducing system. By replacing NADH with NADPH at the same concentration, the efficiecy of AD1 and AD2 were reduced by 70% and 16%, respectively, with dodecanal as the substrate. We also evaluated the kinetic parameters K_M and k_{cat} of AD1 and AD2 in the presence of PMS/NADPH, BSA and dodecanal substrate. From the MM equation, K_M and k_{cat} for AD1 were determined to be 14±8.8 μM and 1.7±0.4/h, respectively; those for AD2 were 11±3.3 μM and 1.7±0.1/h, respectively (Figure 22).

While the k_{cat}/K_M values of AD1 and AD2 were lower when NADPH was used, the turnover numbers of both enzymes were increased by approximately 1.2- to 1.7-fold, indicating that NADPH is able to support the decarbonylation reaction. Both NADH and NADPH are important electron donors in the cell. As the phosphate group in NADPH is distantly located from the electron transfer site, it is unlikely that this phosphate group would affect the reducing capability of NADPH compared to NADH. However, the presence of the phosphate group makes the NADH and NADPH molecules structurally different, thus, the interactions of NADH-cAD and NADPH-cAD can give rise to differing cAD activities.

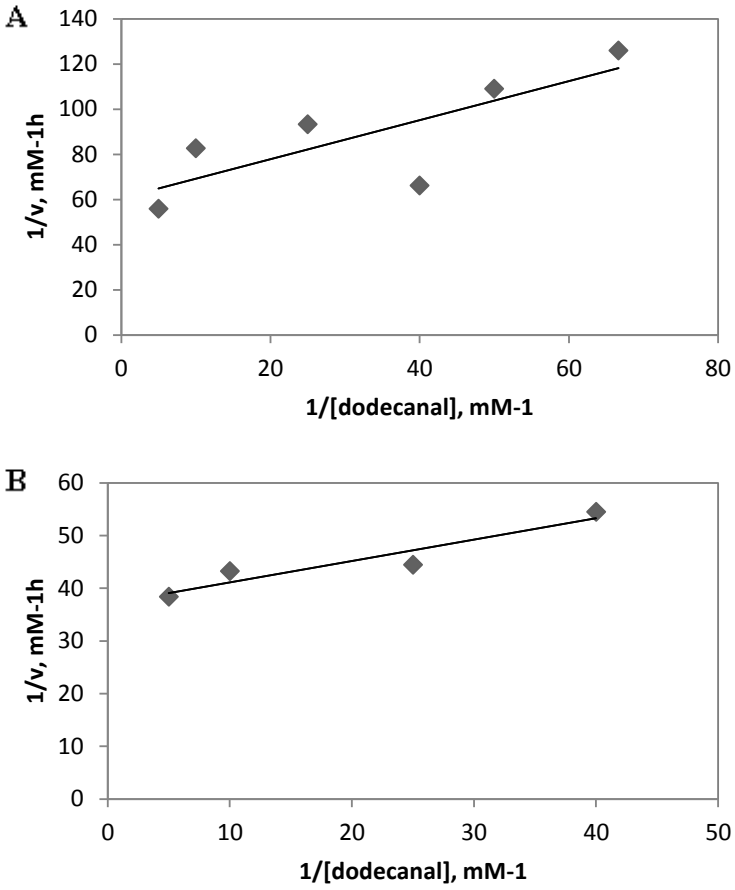


Figure 22 Lineweaver-Burk plots of Michaelis-Menten equation for (A) AD1, with K_M = 14±8.8 µM and k_{cat} = 1.7±0.4/h; and (B) AD2, with K_M = 11±3.3 µM and k_{cat} = 1.7±0.1/h. Reactions contained PMS/NADPH (1 mM), BSA and dodecanal substrate. v is the rate of undecane formation ([undecane]/h). [] indicates concentration.

5. Investigation of *In vivo* decarbonylase activity by feeding aldehyde to the yeast cells

In order to investigate the *in vivo* decarbonylase activity of the ADs from cyanobacteria and insect, these enzymes were recombinantly expressed in *S. cerevisiae*. As aldehyde, the substrate for decarbonylase, does not naturally present in the yeast, external source of various chain length of aldehydes were fed through the growth medium. Following incubation with aldehyde, the cells were then analyzed for hydrocarbon content.

5.1 Cyanobacterial decarbonylase

To determine the ability of cyanobacterial decarbonylase to produce alkane inside yeast cell, *S. cerevisiae* expressing recombinant AD1 were grown in the presence of 0.01% dodecanal (C12 aldehyde) in the growth medium for 24 h, and the yeast strains were harvested and analyzed for alkanes. Wild type *S. cerevisiae* BY4741 was used as a control.

An interesting observation is that when the recombinant strains were grown in the presence of dodecanal, lauric acid, was detected (Figure 23). The acid peaks were very broad and asymmetric as fatty acids are non-volatile and carboxylic groups are very polar and reactive, resulting in inconsistent peak patterns [51]. The presence of dodecanal also inhibits cell growth significantly. After 24 h, cell density was lower by more than 90%, as determined from OD_{600} measurement, compared to when dodecanal was not present. Cell density reduction was observed in the control, as well, with the density of cells grown in the presence of dodecanal was reduced by 75% after incubation.

There could be three reasons for this result; (i) dodecanal is toxic to the yeast, inhibiting the growth significantly and forcing the cells to convert it to the less toxic lauric acid, (ii) the stress-inducible aldehyde dehydrogenase (*ALD*) pathway [52] which converts fatty aldehydes to fatty acids has a higher association constant for the fatty aldehyde substrate compared to cADs and (iii) alkane toxicity towards yeast cells, thereby suppressing cAD activity and hence, alkane formation. Furthermore, the accumulation of lauric acid was only found inside the cells, and not in the medium as analyzed by GC-MS (Figure 23). As lauric acid is not endogenously found in the intracellular compartment of *S. cerevisiae*, the appearance of lauric acid post-dodecanal feeding indicates that the dodecanal was oxidised to lauric acid inside the cell.

In order to verify the inhibitory effect of medium-chain aldehydes on the growth of *S. cerevisiae*, the cells were grown in the presence of three different dodecanal and tetradecanal concentrations, 50 µM, 200 µM and 1000 µM, for prolonged period (3 days). Cells expressing AD1 grew slower than the wild type in general, even when the wild type were grown in the presence of aldehyde, as indicated by longer lag phase of AD1-expressing cells compared to the wild type (Figure 24). This is caused by the metabolic burden exerted by the recombinant expression of AD1. AD1-expressing cells fed with 50 µM, 200 µM or no aldehyde (both dodecanal and tetradecanal experiments exhibited similar profile) appeared to have similar growth profile with the wild type: a significant increase in cell density after 2 days. Cells fed with 1000 µM aldehyde did not grow at all. This could be due to (i) most of the dodecanal and tetradecanal that were present at low concentrations (50 and 200 µM) have already evaporated after 2 days, (ii) the high concentration of the aldehydes (1000 µM) is beyond a certain threshold level of toxicity and hence inhibits growth, or (iii) a combination of both.

Analysis of hydrocarbons extracted from the cells did not indicate the presence of alkanes (possible undecane from dodecanal, or tridecane from tetradecanal). Instead, similar to previous result, lauric and myristic acids were detected when cells were fed with dodecanal and tetradecanal (Appendix C1), respectively, indicating the intracellular oxidation of aldehydes to acids.

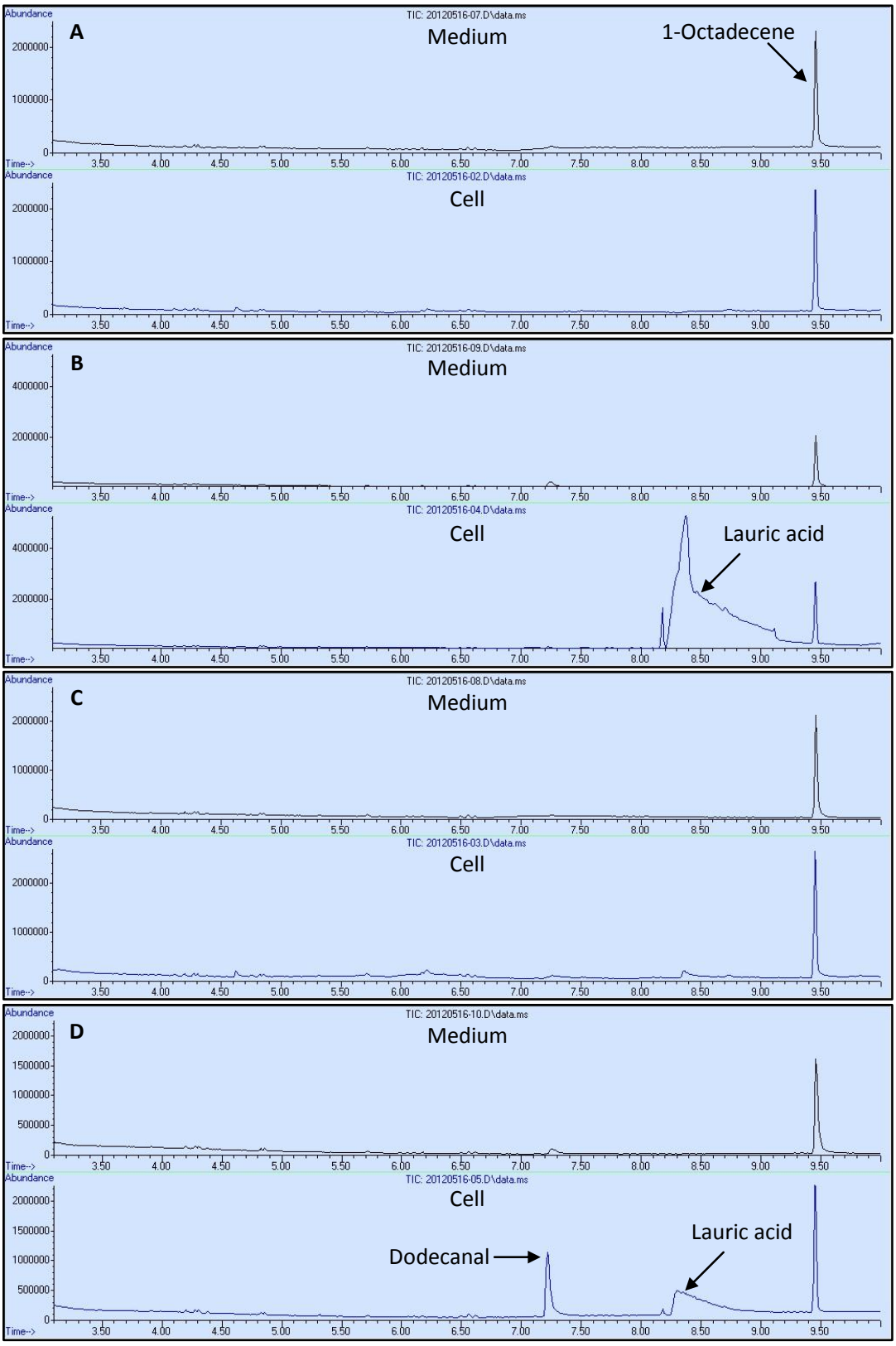


Figure 23 Gas chromatograms of hydrocarbons from (A) control, no aldehyde; (B) control fed with dodecanal; (C) AD1, no aldehyde; (D) AD1 fed with dodecanal. Arrows indicate lauric acid peaks. Peaks on the extreme right side of the chromatographs belong to 1-octadecene, an internal standard.

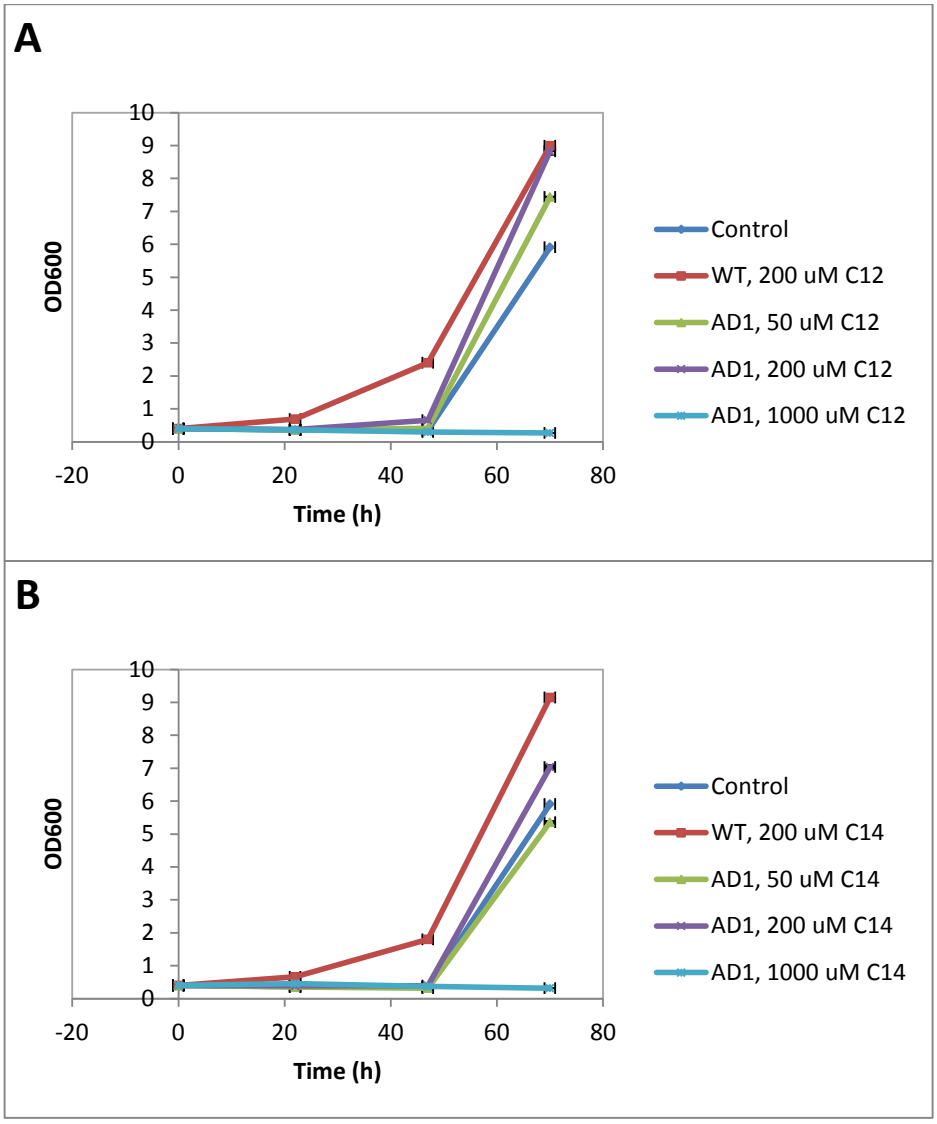


Figure 24 *S. cerevisiae* growth in the presence of (A) dodecanal (C12) and (B) tetradecanal (C14), as measured by cell density (OD600, optical density at 600 nm). Control is AD1-expressing cells grown without aldehyde and WT is wild type *S. cerevisiae*.

In an attempt to promote aldehyde conversion to alkane by increasing enzyme concentration, a whole-cell biocatalytic reaction was performed. In this setup, the whole yeast cell acts as a catalyst for the decarbonylation reaction and the aldehyde substrate can be added to the whole-cell biocatalyst in a reaction buffer. The advantage is that enzyme concentration can be increased simply by increasing cell density in the reaction mixture without the need to isolate and purify the enzyme itself. Such application has been

demonstrated to work in *S. cerevisiae*, for example in the production of l-malic acid [53, 54] and various other products [55-57].

In this study, 25 mg/ml wet cell wt. was incubated with 0.1 mmole of either dodecanal or tetradecanal in 5 ml sodium phosphate buffer (0.1 M, pH 7.2). The buffer was supplemented with $(NH_4)_2Fe(SO_4)_2$ as iron source and 2% dextrose as carbon source. Samples were taken after 2 h and overnight incubation for hydrocarbon analysis. However, intracellular conversion of aldehyde to alkane in yeast still proved challenging, as no alkanes could be detected in the reaction.

5.2 Insect decarbonylase

As the P450 decarbonylase from Musca domestica (CYP4G2) showed oxidative decarbonylase activity when recombinantly expressed in Sf9 cells, here we attempt to express the fusion protein of CYP4G2 and its redox partner NADPH-cytochrome P450 reductase (G2-CPR) in *S. cerevisiae*.

The gene was successfully cloned into the yeast, as verified by DNA analysis of transformed colonies (Figure 25). Expression of recombinant G2-CPR was induced by growing the cells 2% Galactose and 200 µM of dodecanal was included in the medium under two different conditions: (i) immediately added at the start of incubation and (ii) the dodecanal was added one day after incubation. The cultures were grown for a total of three days, before harvested for extraction and analysis of hydrocarbons. In the control, aldehyde was not present.

The growth profile was similar to those observed previously. In condition (i), the yeast did not grow for the first two days, after which the cell density started to increase, indicating the evaporation of the toxic aldehyde (Appendix C2). In condition (ii), the yeast grew normally

until the addition of dodecanal at t = 1 day, after which no significant increase in cell density was observed (Figure 26). To avoid complications from aldehyde evaporation, the experiment was repeated using longer chain aldehydes, hexadecanal and octadecanal. These subtrates are less water-soluble and much less volatile compared to dodecanal. However, based on analysis by GC-MS, alkanes were not detected in all cases.

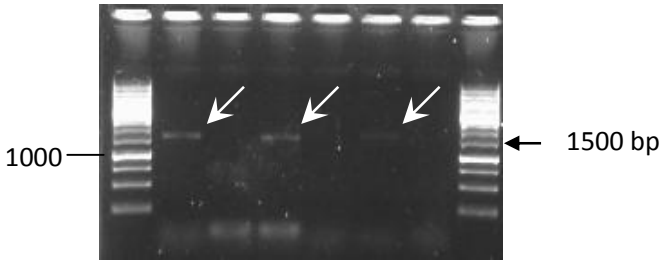


Figure 25 DNA analysis image of G2-CPR gene in *S. cerevisiae*. White arrows indicate positive fragments (approx. 1500 bp).

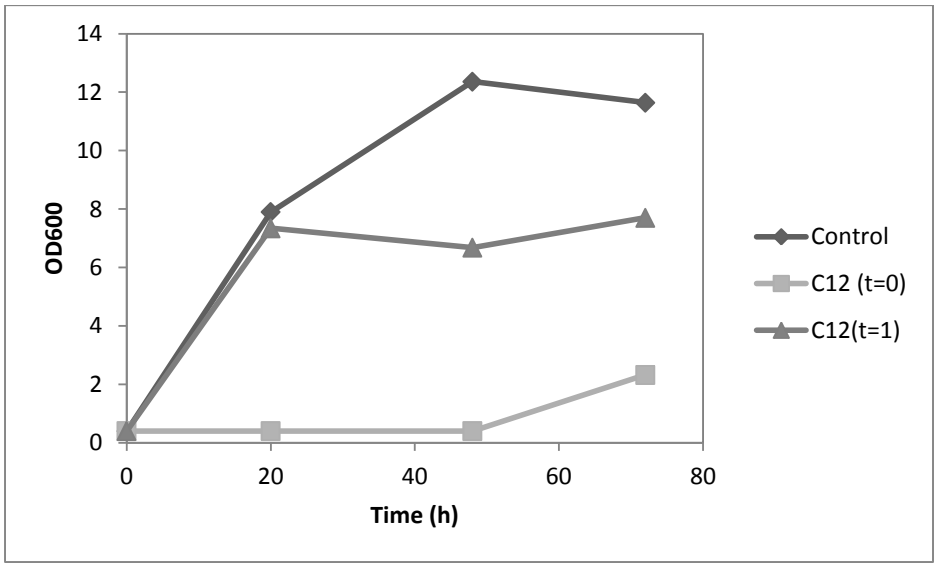


Figure 26 Growth of *S. cerevisiae* expressing G2-CPR in the presence of dodecanal: (i) dodecanal was added immediately at the start of incubation (C12(t=0)) and (ii) dodecanal added one day after the start of incubation (C12(t=1)). Control is cells grown without aldehyde.

5.3 Small Ubiquitin-like Modifier protein

Small Ubiquitin-like Modifier (SUMO) proteins are a class of small ubiuitin-like proteins that are attached to various other proteins *in vivo* as part of a post-translational modification in eukaryotes. This modification often has positive effects on proteins it is attached to, such as improving substrate interactions with protein molecules, promoting protein-protein interactions and assembly of multi-protein complexes [58]. Sumoylation also induces protein targeting to specific cellular compartments or structures and is involved in regulating certain pathways and biological processes such as translation [59]. In yeast, there is only a single gene responsible for the expression of SUMO, *Smt3* [60].

In this study, fusion proteins of SUMO and cyanobacterial decarbonylases SUMO-AD1, SUMO-AD2 and SUMO-AD3 were constructed in an attempt to promote the correct folding of the decarbonylases and their soluble expression in *S. cerevisiae*. The SUMO modification might also improve protein-substrate interaction between decarbonylase and aldehyde substrate.

The three fusion constructs were successfully cloned into *S. cerevisiae* (Figure 27), and the *in vivo* decarbonylation activity was tested using these modified enzymes. The addition of SUMO did not appear to affect the growth of yeast (in comparison to cells expressing non-fusion AD1, AD2 and AD3). After expressions of SUMO-AD1, SUMO-AD2 and SUMO-AD3 in 2% Galactose for one day, 400 µM of octadecanal was added to the cultures. The cells were incubated for another one day before harvested for hydrocarbon extraction and analysis. The results, however, suggested that the SUMO might not have positive effects on the decarbonylase activity of AD1, AD2 and AD3, as no alkanes were detected.

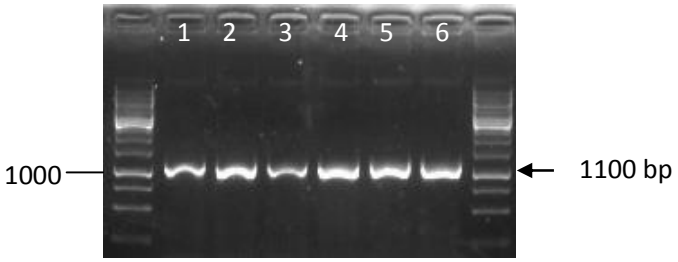


Figure 27 DNA analysis image of SUMO-AD1 (1,2), SUMO-AD2 (3,4) and SUMO AD3 (5,6) genes. All colonies have positive fragments (approx. 1100 bp).

6. Conclusions and Future Works

In conclusion, three cADs from *A. variabilis* (AD1), *A. platensis* (AD2) and *S. elongatus* (AD3), respectively, have been successfully cloned into *S. cerevisiae*, expressed and purified. These cADs exhibited soluble expression in *S. cerevisiae* and *in vitro* enzyme assays indicated that they are biologically active. Characterization of enzyme kinetics showed that AD1 and AD2 show activity, though very low, towards dodecanal substrate; with K_M = 3.5±2.9 µM and k_{cat} = 1.4±0.1/h for AD1 and 5.0±2.2 µM and 0.9±0.1/h for AD2, respectively, in the presence of NADH as a reducing cofactor. NADPH also showed the ability to support decarbonylation reaction albeit with less efficiency.

On the other hand, the *in vivo* conversion of aldehyde to alkane proved to be difficult. Using the recombinantly expressed cyanobacterial and insect decarbonylases, *S. cerevisiae* has not been able to produce alkane on its own thus far. Cyanobacterial AD1, AD2, AD3 and their fusion counterparts SUMO-AD1, SUMO-AD2, SUMO-AD3, and insect P450 decarbonylase G2-CPR have not shown *in vivo* activity towards a range of substrate: dodecanal, tetradecanal, hexadencanal and octadecanal. There could be several factors affecting the *in vivo* activity of the decarbonylase that can be improved, such as the intracellular environment of *S. cerevisiae* and certain metabolic pathways that compete with alkane production pathway.

Based on our preliminary results, further investigations to improve the performance of decarbonylase enzyme, such as protein engineering to broaden substrate specificity for production of various hydrocarbons, will also pose an interesting answer to microbial production of biofuels.

6.1 cAD specificity towards different substrates

The results of our study show that while cADs are active *in vitro* for the dodecanal substrate, no alkanes were detected when decanal was incubated with AD1 in aerobic conditions. Specificity towards different fatty aldehyde substrates with different carbon chain length is an important parameter to facilitate the production of medium carbon chain length alkanes. As cyanobacteria are known to produce longer chain alkanes such as heptadecane [46], it is likely that the AD1 and AD2 enzymes obtained from cyanobacteria have higher activity towards longer chain substrates that are naturally present in their native organisms. Hence, protein engineering might be required to engineer recombinant cADs that have improved activities towards medium chain aldehydes. The protein engineering approach that we intend to use is described in Section 6.1.1.

6.1.1 Rational design of proteins

As the crystal structure of a PCC7942_1593 (AD3) ortholog from P. marinus MIT9313, PMT1231, has been solved (PDB 2OC5) [1], a rational protein design could be used to engineer cADs. The X-ray structure confirmed that PMT1231 is a ferritin-like metal binding protein with a di-iron centre and metal coordinating ligands at its active site. Rational protein design by site-directed mutagenesis is very effective in engineering enzymes to improve its desired activity or selectivity. The downside of this method is that it requires intensive knowledge and information on protein sequence, structure, and function/mechanism. However, it is possible to increase stability, activity or selectivity of enzymes whose structural data are unavailable using their sequence information and molecular modelling of homologous enzymes [61].

6.2 Yeast metabolic pathway engineering

As discussed in Section 2.4, there are two main pathways in *S. cerevisiae* that will possibly interfere with the alkane production pathway by metabolizing aldehyde, the substrate for decarbonylase. These two pathways are aldehyde oxidation to fatty acid and aldehyde reduction to alcohol.

6.2.1 Aldehyde dehydrogenase

We hypothesize that aldehydes may confer toxicity towards yeast cells and hence increase the driving force for conversion of fatty aldehydes to fatty acids by *ALD* gene products. This scenario is undesirable as fatty aldehydes are alkane precursors in the proposed pathway. This pathway will compete with the AD pathway for aldehyde substrate, inhibiting alkane production *in vivo*. Hence, the deletion of *ALD* genes in *S. cerevisiae* could have a significant impact on the *in vivo* activity of ADs. White *et al*. [62] reported that at least two genes, *ALD2* and *ALD3*, are responsible for the conversion of aldehyde (3-aminopropanal) to a carboxylic acid (β-alanine) in *S. cerevisiae*.

6.2.2 Aldehyde reductase

Some alcohol dehydrogenases that are present in *S. cerevisiae* are known to act as aldehyde reductases, reducing short and medium chain aldehydes to alcohol [40, 42]. As such, *in vivo* decarbonylation study using *ΔAdh* mutants in yeast may prove useful. In this study, *ΔAld2* single and *ΔAld2ΔAdh7* double mutants were constructed. Deletion of other *ALD*s and *ADH*s, and a combination of multiple deletions are likely to be important to eliminate competing pathways.

6.2.3 Other important pathways

Other important pathways that are worth studying are the fatty acid synthesis and beta oxidation pathways. Alteration of these pathways may modify the intracellular fatty acid pool. Deletion of beta oxidation genes might hinder fatty acid breakdown, hence increasing free fatty acid levels. Pathways that generate NADPH are also essential, as NADPH is an important coenzyme that is required for many redox reactions inside the cell. Cellular levels of NADPH could also be artificially increased by engineering the malic enzyme metabolic pathway [50] or the ammonium assimilation pathway [63] to facilitate cAD activity.

6.3 Yeast intracellular condition

One possible reason that ADs did not show any activity *in vivo* is because the intracellular condition of *S. cerevisiae* does not support the decarbonylation reaction. Intracellular pH plays a key role in determining the optimal reaction condition since protein solubility, and thus activity, greatly depends on it. Yeast intracellular pH is known to be generally low: around 5.7 during lag phase, and increasing slightly during exponential growth and decreasing to 5.5 during stationary phase [64]. The pH is even lower in the presence of glucose [65], which is the carbon source used in this study. This means that the intracellular pH differs quite significantly from the pH used in *in vitro* study (neutral pH), where alkane formation were observed, even though very slow. The use of growth media with high (external) pH is therefore recommended to promote higher intracellular pH. The other factor is the amount of Fe(II) present in yeast, as decarbonylase enzymes are iron-dependent and iron is part of its catalytic centre. In the YNB minimal media that was used in this study, iron is present only in trace amount. Although Fe(II) in the form of $(NH_4)_2Fe(SO_4)_2$ was

supplemented to the growth media for *in vivo* study, optimal concentration and types of Fe ion required remains to be investigated.

6.4 Co-expression of FAR and AD

Eventually, as the main objective of this study is to produce alkanes from fatty acids in *S. cerevisiae*, the FAR and AD pathways eventually need to be streamlined. To achieve this, a co-expression system using yeast plasmid will be synthesized for *in vivo* characterization of the enzymatic activity of cADs before integrating the two pathways into the genome. To further increase the driving force for alkane production, the β-oxidation-deletion mutants (*pox1*Δ, *pox2*Δ, *pox3*Δ single, double and triple knockouts) will also be used in this study.

References

1. Schirmer, A., et al., *Microbial Biosynthesis of Alkanes.* Science, 2010. **329**(5991): p. 559-562.
2. Aarts, M.G., et al., *Molecular characterization of the CER1 gene of arabidopsis involved in epicuticular wax biosynthesis and pollen fertility.* The Plant Cell Online, 1995. **7**(12): p. 2115-27.
3. Hannoufa, A., J. McNevin, and B. Lemieux, *Epicuticular waxes of eceriferum mutants of Arabidopsis thaliana.* Phytochemistry, 1993. **33**(4): p. 851-855.
4. Sturaro, M., et al., *Cloning and Characterization of GLOSSY1, a Maize Gene Involved in Cuticle Membrane and Wax Production.* Plant Physiology, 2005. **138**(1): p. 478-489.
5. Kunst, L. and A.L. Samuels, *Biosynthesis and secretion of plant cuticular wax.* Progress in Lipid Research, 2003. **42**(1): p. 51-80.
6. Yoder, J.A., et al., *Enhancement of diapausing flesh fly puparia with additional hydrocarbons and evidence for alkane biosynthesis by a decarbonylation mechanism.* Insect Biochemistry and Molecular Biology, 1992. **22**(3): p. 237-243.
7. Reed, J.R., et al., *Unusual mechanism of hydrocarbon formation in the housefly: cytochrome P450 converts aldehyde to the sex pheromone component (Z)-9-tricosene and CO2.* Proceedings of the National Academy of Sciences, 1994. **91**(21): p. 10000-10004.
8. Das, D., et al., *Oxygen-Independent Decarbonylation of Aldehydes by Cyanobacterial Aldehyde Decarbonylase: A New Reaction of Diiron Enzymes.* Angewandte Chemie International Edition, 2011. **50**(31): p. 7148-7152.
9. Eser, B.E., et al., *Oxygen-Independent Alkane Formation by Non-Heme Iron-Dependent Cyanobacterial Aldehyde Decarbonylase: Investigation of Kinetics and Requirement for an External Electron Donor.* Biochemistry, 2011. **50**(49): p. 10743-10750.
10. Dennis, M.W. and P.E. Kolattukudy, *Alkane biosynthesis by decarbonylation of aldehyde catalyzed by a microsomal preparation from Botryococcus braunii.* Archives of Biochemistry and Biophysics, 1991. **287**(2): p. 268-275.
11. Cheesbrough, T.M. and P.E. Kolattukudy, *Alkane biosynthesis by decarbonylation of aldehydes catalyzed by a particulate preparation from Pisum sativum.* Proceedings of the National Academy of Sciences, 1984. **81**(21): p. 6613-6617.
12. Richardson, A., et al., *Cloning and expression analysis of candidate genes involved in wax deposition along the growing barley (<i>Hordeum vulgare</i>) leaf.* Planta, 2007. **226**(6): p. 1459-1473.
13. Hu, X., et al., *cDNA cloning and expression analysis of a putative decarbonylase <i>TaCer1</i> from wheat (<i>Triticum aestivum</i> L.).* Acta Physiologiae Plantarum, 2009. **31**(6): p. 1111-1118.
14. Qiu, Y., et al., *An insect-specific P450 oxidative decarbonylase for cuticular hydrocarbon biosynthesis.* Proceedings of the National Academy of Sciences, 2012. **109**(37): p. 14858-14863.
15. Chung, H., et al., *Characterization of Drosophila melanogaster cytochrome P450 genes.* Proceedings of the National Academy of Sciences, 2009.

16. Fan, Y., et al., *Hydrocarbon synthesis by enzymatically dissociated oenocytes of the abdominal integument of the German Cockroach, Blattella germanica.* Naturwissenschaften, 2003. **90**(3): p. 121-126.
17. Reed, J.R., et al., *Proposed mechanism for the cytochrome P 450-catalyzed conversion of aldehydes to hydrocarbons in the house fly, Musca domestica.* Biochemistry, 1995. **34**(49): p. 16221-16227.
18. Müller, J., et al., *NMR Structure of the [2Fe-2S] Ferredoxin Domain from Soluble Methane Monooxygenase Reductase and Interaction with Its Hydroxylase†,‡.* Biochemistry, 2001. **41**(1): p. 42-51.
19. Wu, C.-H., et al., *YfaE, a Ferredoxin Involved in Diferric-Tyrosyl Radical Maintenance in Escherichia coli Ribonucleotide Reductase†.* Biochemistry, 2007. **46**(41): p. 11577-11588.
20. Warui, D.M., et al., *Detection of Formate, Rather than Carbon Monoxide, As the Stoichiometric Coproduct in Conversion of Fatty Aldehydes to Alkanes by a Cyanobacterial Aldehyde Decarbonylase.* Journal of the American Chemical Society, 2011. **133**(10): p. 3316-3319.
21. Halaka, F.G., G.T. Babcock, and J.L. Dye, *Properties of 5-methylphenazinium methyl sulfate. Reaction of the oxidized form with NADH and of the reduced form with oxygen.* Journal of Biological Chemistry, 1982. **257**(3): p. 1458-1461.
22. NAKANO, K., et al., *Application of the Enzymic Electric Cell Method to the Activity Assay of NAD-linked Dehydrogenases.* Journal of Biochemistry, 1975. **78**(6): p. 1347-1352.
23. Anderson, R.F., *Energetics of the one-electron steps in the NAD+/NADH redox couple.* Biochimica et Biophysica Acta (BBA) - Bioenergetics, 1980. **590**(2): p. 277-281.
24. Farrington, J.A., E.J. Land, and A.J. Swallow, *The one-electron reduction potentials of NAD.* Biochimica et Biophysica Acta (BBA) - Bioenergetics, 1980. **590**(2): p. 273-276.
25. Spector, A.A., K. John, and J.E. Fletcher, *Binding of long-chain fatty acids to bovine serum albumin.* Journal of Lipid Research, 1969. **10**(1): p. 56-67.
26. Charbonneau, D.M. and H.-A. Tajmir-Riahi, *Study on the Interaction of Cationic Lipids with Bovine Serum Albumin†.* The Journal of Physical Chemistry B, 2009. **114**(2): p. 1148-1155.
27. Li, N., et al., *Conversion of Fatty Aldehydes to Alka(e)nes and Formate by a Cyanobacterial Aldehyde Decarbonylase: Cryptic Redox by an Unusual Dimetal Oxygenase.* Journal of the American Chemical Society, 2011. **133**(16): p. 6158-6161.
28. Eser, B.E., et al., *Correction to Oxygen-Independent Alkane Formation by Non-Heme Iron-Dependent Cyanobacterial Aldehyde Decarbonylase: Investigation of Kinetics and Requirement for an External Electron Donor.* Biochemistry, 2012.
29. Li, N., et al., *Evidence for Only Oxygenative Cleavage of Aldehydes to Alk(a/e)nes and Formate by Cyanobacterial Aldehyde Decarbonylases.* Biochemistry, 2012. **51**(40): p. 7908-7916.
30. Andre, C., et al., *Fusing catalase to an alkane-producing enzyme maintains enzymatic activity by converting the inhibitory byproduct H2O2 to the cosubstrate O2.* Proceedings of the National Academy of Sciences, 2013. **110**(8): p. 3191-3196.
31. Verma, R., E. Boleti, and A.J.T. George, *Antibody engineering: Comparison of bacterial, yeast, insect and mammalian expression systems.* Journal of Immunological Methods, 1998. **216**(1–2): p. 165-181.
32. Kapust, R.B. and D.S. Waugh, *Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused.* Protein Science, 1999. **8**(8): p. 1668-1674.
33. Romanos, M.A., C.A. Scorer, and J.J. Clare, *Foreign gene expression in yeast: a review.* Yeast, 1992. **8**(6): p. 423-488.

34. Fickers, P., et al., *Hydrophobic substrate utilisation by the yeast Yarrowia lipolytica, and its potential applications.* FEMS Yeast Research, 2005. **5**(6–7): p. 527-543.
35. Lasserre, J.-P., et al., *First complexomic study of alkane-binding protein complexes in the yeast Yarrowia lipolytica.* Talanta, 2010. **80**(4): p. 1576-1585.
36. Tehlivets, O., K. Scheuringer, and S.D. Kohlwein, *Fatty acid synthesis and elongation in yeast.* Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids, 2007. **1771**(3): p. 255-270.
37. Hiltunen, J.K., et al., *Peroxisomal multifunctional beta-oxidation protein of Saccharomyces cerevisiae. Molecular analysis of the fox2 gene and gene product.* Journal of Biological Chemistry, 1992. **267**(10): p. 6646-6653.
38. Hiltunen, J.K., et al., *The biochemistry of peroxisomal β-oxidation in the yeast Saccharomyces cerevisiae.* FEMS Microbiology Reviews, 2003. **27**(1): p. 35-64.
39. Grey, M., M. Schmidt, and M. Brendel, *Overexpression of ADH1 confers hyper-resistance to formaldehyde in Saccharomyces cerevisiae.* Current Genetics, 1996. **29**(5): p. 437-440.
40. Larroy, C., et al., *Characterization of the Saccharomyces cerevisiae YMR318C (ADH6) gene product as a broad specificity NADPH-dependent alcohol dehydrogenase: relevance in aldehyde reduction.* Biochem. J., 2002. **361**(1): p. 163-172.
41. Lewis Liu, Z., et al., *Multiple gene-mediated NAD(P)H-dependent aldehyde reduction is a mechanism of in situ detoxification of furfural and 5-hydroxymethylfurfural by Saccharomyces cerevisiae.* Applied Microbiology and Biotechnology, 2008. **81**(4): p. 743-753.
42. Larroy, C., X. Parés, and J.A. Biosca, *Characterization of a Saccharomyces cerevisiae NADP(H)-dependent alcohol dehydrogenase (ADHVII), a member of the cinnamyl alcohol dehydrogenase family.* European Journal of Biochemistry, 2002. **269**(22): p. 5738-5745.
43. Dickinson, F.M., *The purification and some properties of the Mg(2+)-activated cytosolic aldehyde dehydrogenase of Saccharomyces cerevisiae.* Biochem. J., 1996. **315**(2): p. 393-399.
44. Meaden, P.G., et al., *The ALD6 gene of Saccharomyces cerevisiae encodes a cytosolic, Mg2+-activated acetaldehyde dehydrogenase.* Yeast, 1997. **13**(14): p. 1319-1327.
45. Jacobson, M.K. and C. Bernofsky, *Mitochondrial acetaldehyde dehydrogenase from Saccharomyces cerevisiae.* Biochimica et Biophysica Acta (BBA) - Enzymology, 1974. **350**(2): p. 277-291.
46. Fehler, S.W.G. and R.J. Light, *Biosynthesis of hydrocarbons in Anabaena variabilis. Incorporation of [methyl-14C] and [methyl-2H2] Methionine into 7- and 8-methyl-heptadecanes.* Biochemistry, 1970. **9**(2): p. 418-422.
47. Arnold, T. and D. Linke, *Phase separation in the isolation and purification of membrane proteins.* Biotechniques, 2007. **43**(4): p. 427-440.
48. Alfalah, M., et al., *A Novel Type of Detergent-resistant Membranes May Contribute to an Early Protein Sorting Event in Epithelial Cells.* Journal of Biological Chemistry, 2005. **280**(52): p. 42636-42643.
49. Alberts, B., et al., *Catalysis and the Use of Energy by Cells*, in *Molecular Biology of the Cell*, Dries and D. J, Editors. 2002, Garland Science: New York.
50. Moreira dos Santos, M., et al., *Manipulation of malic enzyme in Saccharomyces cerevisiae for increasing NADPH production capacity aerobically in different cellular compartments.* Metabolic Engineering, 2004. **6**(4): p. 352-363.
51. Igor, Z., *Acids*, in *Encyclopedia of Chromatography, Third Edition (Print Version)*. 2009, CRC Press.
52. Aranda, A. and M.l. del Olmo, *Response to acetaldehyde stress in the yeast Saccharomyces cerevisiae involves a strain-dependent regulation of several ALD*

genes and is mediatedby the general stress response pathway. Yeast, 2003. **20**(8): p. 747-759.
53. Stojkovič, G. and P. Žnidaršič-Plazl, *Continuous synthesis of l-malic acid using whole-cell microreactor.* Process Biochemistry, 2012. **47**(7): p. 1102-1107.
54. Neufeld, R.J., et al., *l-Malic acid formation by immobilized Saccharomyces cerevisiae amplified for fumarase.* Enzyme and Microbial Technology, 1991. **13**(12): p. 991-996.
55. Kratzer, R., S. Egger, and B. Nidetzky, *Integration of enzyme, strain and reaction engineering to overcome limitations of baker's yeast in the asymmetric reduction of α-keto esters.* Biotechnology and Bioengineering, 2008. **101**(5): p. 1094-1101.
56. Jeong, M., et al., *Optimization of enantioselective synthesis of methyl (R)-2-chloromandelate by whole cells of Saccharomyces cerevisiae.* Biotechnology Letters, 2010. **32**(10): p. 1529-1531.
57. Johanson, T., et al., *Reaction and strain engineering for improved stereo-selective whole-cell reduction of a bicyclic diketone.* Applied Microbiology and Biotechnology, 2008. **77**(5): p. 1111-1118.
58. Johnson, E.S., *PROTEIN MODIFICATION BY SUMO.* Annual Review of Biochemistry, 2004. **73**(1): p. 355-382.
59. Seeler, J.-S. and A. Dejean, *Nuclear and unclear functions of SUMO.* Nat Rev Mol Cell Biol, 2003. **4**(9): p. 690-699.
60. Hannich, J.T., et al., *Defining the SUMO-modified Proteome by Multiple Approaches in Saccharomyces cerevisiae.* Journal of Biological Chemistry, 2005. **280**(6): p. 4102-4110.
61. Bornscheuer, U.T. and M. Pohl, *Improved biocatalysts by directed evolution and rational protein design.* Current Opinion in Chemical Biology, 2001. **5**(2): p. 137-143.
62. White, W.H., et al., *Specialization of Function Among Aldehyde Dehydrogenases: The ALD2 and ALD3 Genes Are Required for ß-Alanine Biosynthesis in Saccharomyces cerevisiae.* Genetics, 2003. **163**(1): p. 69-77.
63. Moreira dos Santos, M., et al., *Aerobic physiology of redox-engineered Saccharomyces cerevisiae strains modified in the ammonium assimilation for increased NADPH availability.* FEMS Yeast Research, 2003. **4**(1): p. 59-68.
64. Imai, T. and T. Ohno, *Measurement of yeast intracellular pH by image processing and the change it undergoes during growth phase.* Journal of Biotechnology, 1995. **38**(2): p. 165-172.
65. Kotyk, A., *Intracellular pH of baker's yeast.* Folia Microbiologica, 1963. **8**(1): p. 27-31.

Appendices

A. List of Primers

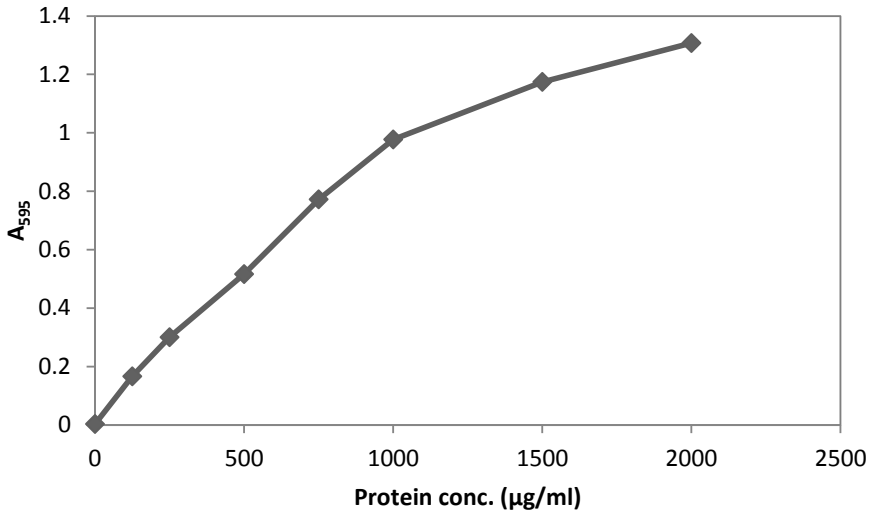
pyes2ct-f: 5’-CCTCTATACTTTAACGTCAAGGAGA-3’

pyes2ct-r: 5’-CTTCCTTTTCGGTTAGAGCGG-3’

cerm-r: 5’- CCGCTCGAGCGGGTGGTGTGGTAACAAC-3’

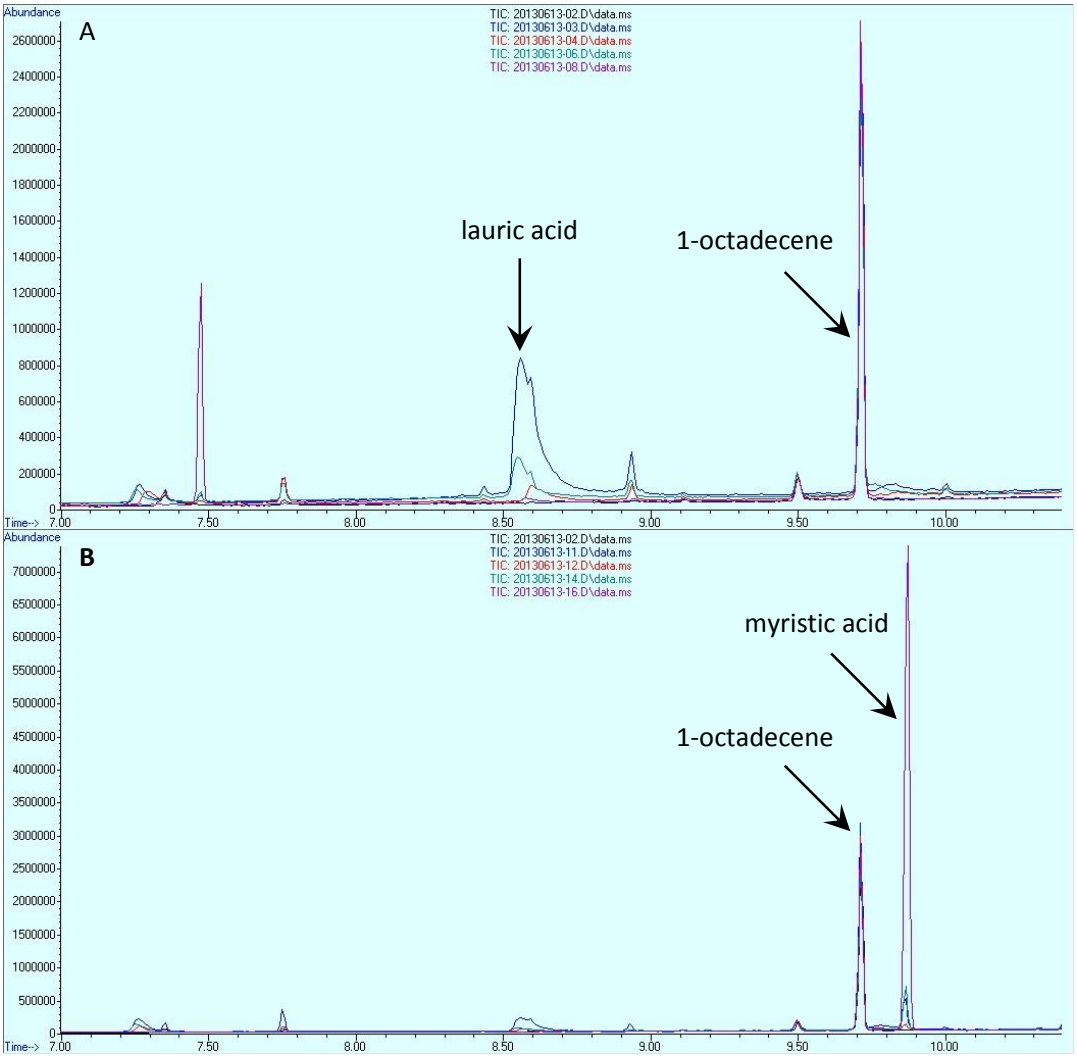
hvcerm-r: 5’- TGAATAACGGAGGTAATAGGTTCAG-3’

B. Standard curve for Bradford assay



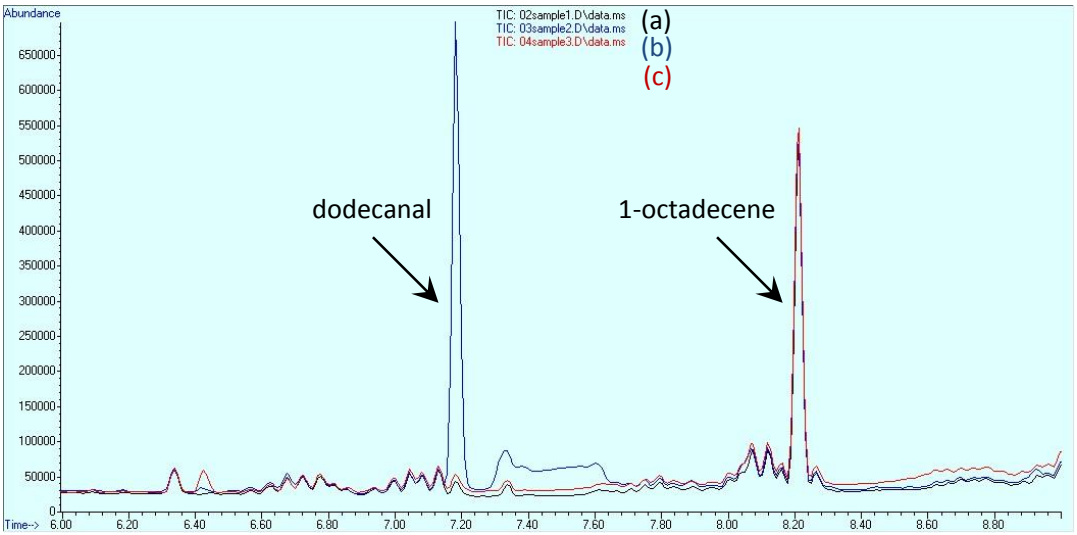
C. Gas chromatogram results of hydrocarbon analysis

C1



Gas chromatogram of hydrocarbons from *S. cerevisiae* expressing AD1 fed with dodecanal (A) and tetradecanal (B). 1-Octadecene was used as an internal standard.

C2



Gas chromatogram of hydrocarbons from *S. cerevisiae* expressing G2-CPR fed with dodecanal and grown for three days. (a) is a control (no aldehyde present), (b) dodecanal was fed at one day after the start of incubation and (c) dodecanal was added immediately at the start of incubation. No more dodecanal was present in (c). 1-Octadecene was used as an internal standard.