Towards the bioproduction of methacrylic acid: a case study on the use of decarboxylases

Luca Rossoni

Thesis submitted to the University of Nottingham
for the degree of Doctor of Philosophy

April 2016

Abstract

In this study, the biocatalytic production of methacrylic acid (MAA), a commodity chemical with a large global market, was investigated. The target was to develop MAA formation *via* enzymatic decarboxylation of itaconic, mesaconic, citraconic and citramalic acid. Several candidate enzymes were tested, but none catalysed the target reactions. Therefore, a novel high throughput screening method for decarboxylases was developed to enable protein engineering. The screening assay detected gaseous CO_2 with a pH-dependent colorimetric reaction, in a 96-well format. Focused mutagenesis was then performed on the mevalonate diphosphate decarboxylases (MVD) from *Saccharomyces cerevisiae* and from *Picrophilus torridus*. However, MAA formation was not observed, even after screening over 3800 variants.

Further investigation was then performed on some of the candidates to understand why MAA was not formed and to develop alternative strategies for enzyme selection. *P. torridus* MVD was discovered to belong to a new class of enzymes, mevalonate-3-kinase, and to be part of a novel archaeal mevalonate pathway. The enzyme also catalysed the formation of isobutene from 3-hydroxyisovalerate with unprecedented production rates. *S. cerevisiae* and *Aspergillus niger* phenylacrylic acid decarboxylases, previously believed to be cofactor-free enzymes, were proposed to require divalent metals and a novel organic cofactor. The enzymatic mechanism of *Pseudomonas fluorescens* α-amino-β-carboxymuconate-ε-semialdehyde (ACMS) decarboxylase was investigated. Through inhibition studies, it was demonstrated that the aldehyde and amino moieties of ACMS could be directly involved in catalysis, in contrast to the previously reported mechanism. This information partially explains why some of these candidates could not catalyse the target reaction, and provides insights into their potential future use for the bioproduction of MAA. Although this route does not seem feasible at present, this study gives foundation and guidance for future investigations.

This thesis is dedicated to my late Grandpa
I hope your Milanese dialect is understood up there

Acknowledgments

My gratitude goes to many people who supported me during the completion of my study.

Thanks to my supervisor, Professor Gill Stephens, for her guidelines and patience in the moments of need, to my co-supervisor, Professor Peter Licence, for his suggestions on chemistry-related issues, and to Graham Eastham and Lucite, for supporting my study even when results were not coming.

Thanks to Steve, for all the good laughs and for becoming the person to freely talk about everything.

Thanks to Andy and Joe, who went with me in this adventure since the very beginning, and to Pawel, who taught me most of the things I needed to know to start my research.

Thanks to Christof, Angela, Lei, Athina, Patricio, Alison, Zoe, Camille, Aidan, Juliana, Alanna, Valentine, Lizeth, Rachel, Anca, Anna and Jaicy, who filled my days at work with fun and entertainment.

Thanks to the Outlaws and to American Football, for allowing me to vent my anger in the right way.

Thanks to Dan, Sara and Mike, now back in US. Without you guys, my life in Nottingham would have been so much less full. You really helped me a lot!

Thanks to my family, for staying with me even when I did not want to, and for showing up in Nottingham so many times in the last four years bringing a lot of Italian "love".

Thanks to all my friends in Italy, the Piazza, especially to Briga and Amptrea. I missed you every day! Thanks for always being there, even from a distance.

Finally, thanks to my beloved Saretta. I know, sometimes it looks like I am not thankful to you at all but, trust me, you are the pillar that holds my life.

Table of contents

List of abbreviations

- $(NH_4)_6Mo_7O_{24}$: ammonium molybdate
- 2-MD: (*R*)-2-methylmalate dehydrogenase
- 3-HAA: 3-hydroxyanthranillic acid
- 3-HB: 3-hydroxybutyrate
- 3-HIV: 3-hydroxyisovaleric acid
- 3-HP: 3-hydroxypropionate
- 3-MA: 3-methlyl aspartase
- 3-PB: 3-phosphobutyrate
- 3-PIV: 3-phosphoisovalerate
- ACH: acetone cyanohydrin
- ACMS: α-amino-β-carboxymuconate-ε-semialdehyde
- ACMSD: *α*-amino-*β*-carboxymuconate-*ε*-semialdehyde decarboxylase
- ACP: acyl carrier protein
- AMS: 2-aminomuconate-6-semialdehyde
- CAD: *cis*-aconitate decarboxylase
- CCT: citramalate-CoA transferase
- CH: citramalate hydrolase
- CS: citramalate synthase
- D_2O: deuterium oxide
- dH_2O: distilled water
- DI: decolouring index
- DMSO: dimethyl sulfoxide
- EDTA: ethylenediaminetetracetic acid
- EHMA: ethylhexyl methacrylic acid
- EMA: ethyl methacrylic acid
- epPCR: error prone PCR
- ESI: electrospray ionization
- $EtBr_2$: ethidium bromide
- EtOH: ethanol
- GC: gas chromatography
- GM: glutamate mutase
- H_2SO_4: sulphuric acid
- HAO: 3,4-hydroxyanthranilinic acid dioxygenase
- HCl: hydrochloric acid
- HCN: hydrogen cyanide
- HPLC: high pressure liquid chromatography
- *i*-BMA: isobutyl methacrylic acid
- IH: itaconyl-CoA hydratase
- IP: isopentenyl phosphate
- IPK: isopentenyl phosphate kinase
- IPP: isopentenyl pyrophosphate
- IPTG: isopropyl β-D-1-thiogalactopyranoside
- ISM: iterative saturation mutagenesis
- KCl: potassium chloride
- $KMnO_4$: potassium permanganate
- LB: : Luria Bertani
- LDH: lactate dehydrogenase
- M3K: mevalonate-3-kinase
- M3P5K: mevalonate 3-phosphate 5-kinase
- M5K: mevalonate-5-kinase
- MAA: methacrylic acid
- MD: malonate decarboxylase
- $MgCl_2$: magnesium chloride
- $MgSO_4$: magnesium sulphate
- MMA: methyl methacrylic acid
- MS: mass spectrometry
- MSA: multiple sequence alignment
- MVA: mevalonate
- MVA-3,5PP: (*R*)-mevalonate-3,5-bisphosphate
- MVA-3P: mevalonate 3-phosphate
- MVA-5P: (*R*,*S*)-mevalonate-5-phosphate
- MVAPP: (*R*,*S*)-mevalonate-5-diphosphate
- MVD: mevalonate diphosphate decarboxylase
- MW: molecular weight
- NaCl: sodium chloride
- $NaHCO_3$: sodium carbonate
- NaOH: sodium hydroxide
- *n*-BMA: n-butyl methacrylic acid
- NMR: nuclear magnetic resonance
- OAD: oxaloacetate decarboxylase
- PAD: phenyl acrylic acid decarboxylase
- PAGE: polyacrylamide gel electrophoresis
- PCR: polymerase chain reaction
- PDC: pyruvate decarboxylase
- PEP: phoshoenolpyruvate
- Pi: inorganic phosphate
- PK: pyruvate kinase
- PLP: pyridoxal-5-phosphate
- PMD: MVA-5P decarboxylases
- PMMA: polymethyl methacrylic acid
- PTFE: polytetrafluoroethylene
- QA: quinolinic acid
- SDS: sodium dodecyl sulphate
- SIM: selected ion monitoring
- ST: sulphotransferase
- TCA: tricarboxylic acid
- TE: thioesterase
- TEMED: tetramethylethylenediamine
- ThDP: thiamine diphosphate
- $ZnCl_2$: zinc chloride

List of figures

List of tables

1 Introduction

Methacrylic acid (MAA) is a commodity chemical with a variety of applications for manufacturing synthetic resins and plastics, electronics, computers, medicine, paints and coatings (Methacrylate Producers Association, 2012). MAA is mainly used for the production of methylmethacrylate (MMA), the monomer for the synthesis of polymethylmethacrylate (PMMA). PMMA is a transparent plastic used as a substitute for glass, which has a global annual demand higher than 3 million metric tons (Information Handling Services, 2014). MMA and MAA are currently manufactured through energy intensive chemical processes which use oil-derived precursors (Bauer, 2000; Yang, 2004). Although these processes are very efficient, concerns regarding their future economics and sustainability have been expressed, and the major MAA industries have shown their interest in the development of a biocatalytic process (de Guzman, 2013; Lucite International, 2014; Mitsubishi Rayon Co., 2011).

MAA is not a naturally occurring chemical, and there is no natural pathway or enzyme that can be used for its bioproduction. Nevertheless, there are some natural metabolic intermediates that could potentially be converted to MAA in a single step reaction, if suitable enzymes can be identified. Among these, itaconic, mesaconic, citraconic and citramalic acid are very attractive, as they could potentially be converted to MAA *via* a simple decarboxylation or decarboxylation/dehydration. This route would require the addition of only one final step to pre-existing pathways and, therefore, it represents an inviting strategy for the direct production of MAA. In this study, the use of decarboxylases to catalyse that step was investigated.

2 Literature review

2.1 Methacrylic acid

Methacrylic acid (MAA) (Figure 2.1), or 2-methyl-2-propenoic acid, is a low molecular weight carboxylic acid with the formula $C_4H_6O_2$, in the form of a corrosive liquid with an acrid unpleasant odour.

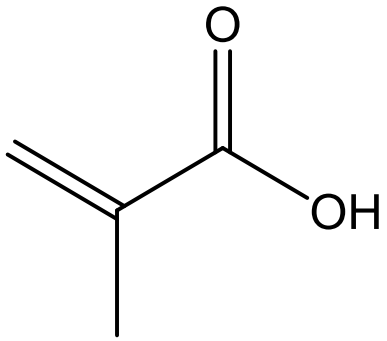


Figure 2.1: Methacrylic acid structure

MAA is used as a building block to make a wide variety of polymers because it is very reactive and can be easily esterified to form numerous derivatives. Examples are *n*-butyl methacrylate (*n*-BMA), isobutyl methacrylate (*i*-BMA), ethylhexyl methacrylate (EHMA), ethyl methacrylate (EMA) and methyl methacrylate (MMA). All these MAA-derived products have several applications in the electronics and computer industries, in biocompatible medical and dental polymers, in paints and coatings and in the production of different synthetic resins (Methacrylate Producers Association, 2012). The most widely used MAA ester is MMA. MMA is mainly used for the production of polymethyl methacrylate (PMMA), a transparent plastic which is used as a substitute for glass. PMMA properties are unique: it has an exceptional optical clarity, a high stability to weathering, a very low UV absorbance, a strong ability to be coloured or moulded and a good bio-compatibility with the human body. Moreover, PMMA can be blended with other chemicals to provide additional properties and it is also recyclable to the virgin monomer (Smolders and Baeyens, 2004; Zhang *et al.*, 2002).

Under the name of two main brands, Lucite and Plexiglas, PMMA is the MAA derivative with the largest market. Its global annual demand has seen a gradual increase of around 4% per year since 2008, and the annual worldwide need is estimated to reach a total of 3.90 million metric tons by 2018, a market worth more than £3 billion (de Guzman, 2013; Information Handling Services, 2014). These numbers only refer to the PMMA market, and values would be even higher with all the other MAA derivatives.

2.1.1 Current production

Several chemical routes for the production of MAA or MMA have been reported, but only a few have achieved commercialisation (Bauer, 2000; Yang, 2004). The oldest production process, known as the ACH route, starts from acetone and hydrogen cyanide (HCN) (Figure 2.2A). HCN and dry acetone react to generate acetone cyanohydrin (ACH), which is then reacted with sulphuric acid in excess to form methacrylamide sulphate. Methacrylamide sulphate is then treated with methanol to form MMA or hydrolysed in the presence of water to form MAA (Bauer, 2000).

The ACH route is very effective and is still the main process used by industry. However, it requires the constant recycling of sulphuric acid and the ability to efficiently dispose of ammonium sulphate (Bauer, 2000), which is released in the last step of the process (Figure 2.2A). Moreover, HCN is very toxic and many difficulties are encountered in its handling and transportation. A slightly different alternative has been developed by Mitsubishi Rayon (Nagai, 2001), where ACH is produced in the same way but then converted to MMA through a different route (Figure 2.2B). This process, known as the MGC route, allows recycling of HCN and avoids the use of sulphuric acid.

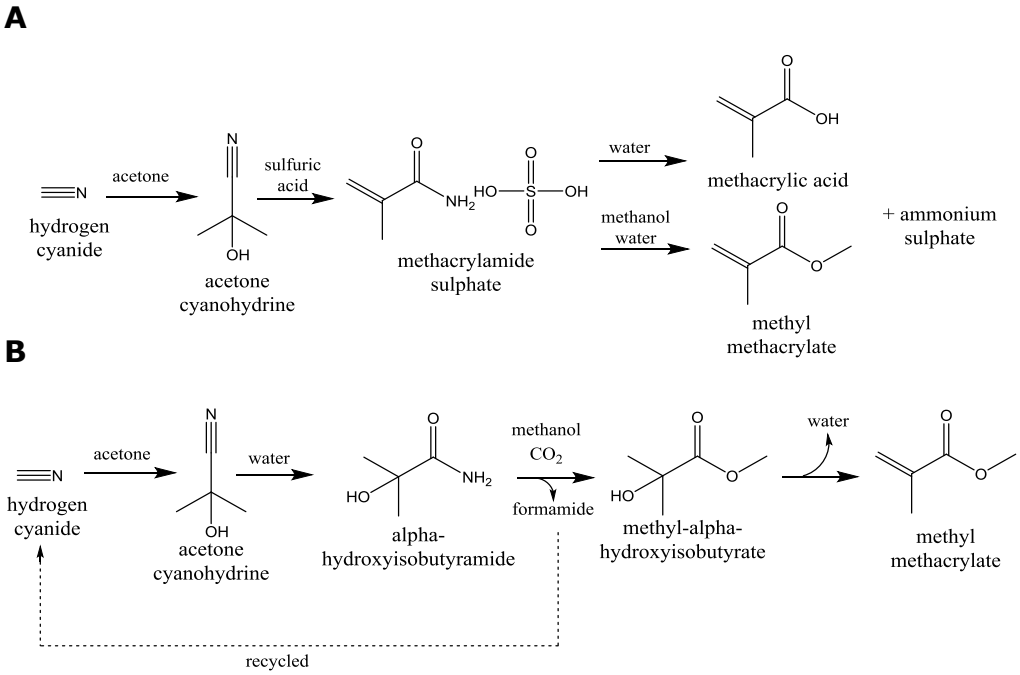
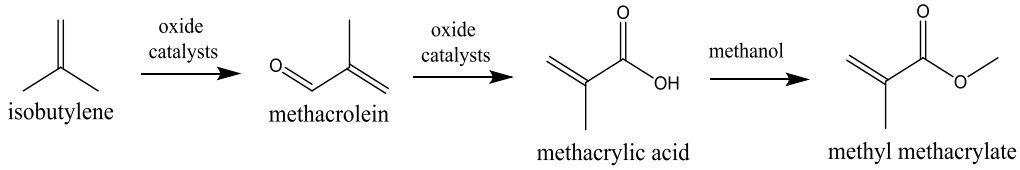


Figure 2.2: ACH and MGC route
(A) The ACH route and (B) the MGC route.

Other technologies commercially employed are the isobutene and isobutyric acid routes (Bauer, 2000). In the isobutene route (Figure 2.3A), isobutene is first oxidised to methacrolein which is further oxidised to MAA in the presence of metal co-catalysts. In the isobutyric acid route (Figure 2.3B), propylene is used as starting material and it is converted to isobutyric acid first in the presence of carbon monoxide, water and a strong acid. Isobutyric acid is then dehydrogenated to MAA with the help of phosphomolybdic acid derivatives (Bauer, 2000). As for the ACH route, MAA can be esterified with methanol to produce MMA. Although these technologies have the advantage of avoiding the use of HCN, they still rely on the utilisation of strong acids and/or energy intensive reaction conditions.

A



B

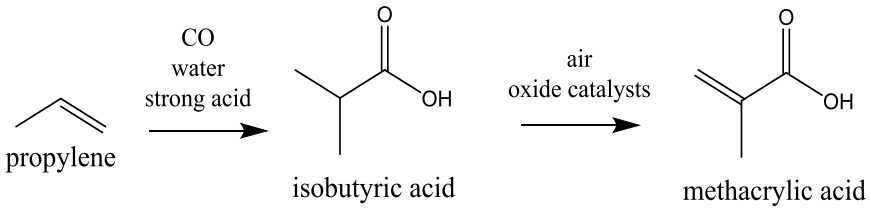


Figure 2.3: Isobutene and isobutyric acid route
(A) The isobutene route and (B) the isobutyric acid route.

A more recent process for production of MMA limits some of the problems mentioned above (Figure 2.4) (Harris, 2010). Known as the "Alpha process", it consists of two main reactions. In the first reaction, carbon monoxide, methanol and ethylene are converted to methyl propionate with no by-products. In the second reaction, methyl propionate is reacted with formaldehyde to form MMA and water under almost anhydrous conditions. No significant hazardous chemicals and little waste treatment are required, by-product formation is low and raw material efficiency exceeds all other MMA technologies.

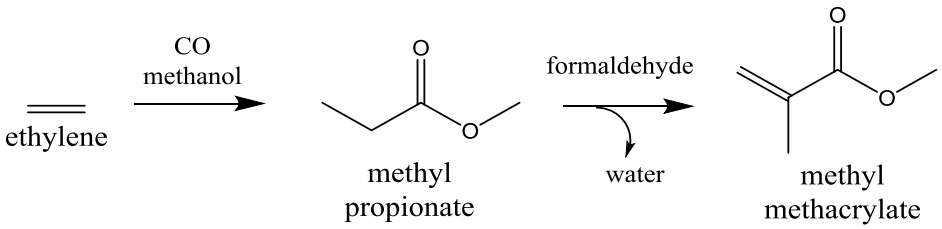


Figure 2.4: Alpha process

2.1.2 Towards bioproduction

The industrial production processes for MAA and MMA are all very energy intensive and rely on petrochemicals as starting materials. The price of crude oil fluctuates drastically, which reflects even more

dramatically on the price of its derivatives (InvestmentMine, 2015). This is a problem, because it means that the petrochemical economy is characterised by unpredictable costs, which can negatively affect supply and demand (Lucite International, 2012). For this reason, the major MAA and MMA industries have expressed their interest in developing alternatives (de Guzman, 2013; Lucite International, 2014; Mitsubishi Rayon Co., 2011). Moreover, there is an increasing public awareness of the environmental impact of oil extraction and of a possible raw oil shortage, which has driven political leaders to promote the utilisation of more sustainable feedstock. In addition, the global concerns over environmental pollution and climate change through the effect of greenhouse gases have pushed the petrochemical industries to search for novel ways to reduce carbon footprint (Becker and Wittmann, 2015). All these concerns could be addressed with the development of a bio-based alternative for MAA or MMA production, through synthetic biology and biocatalysis (de Guzman, 2013; Lucite International, 2014; Mitsubishi Rayon Co., 2011).

Biocatalysis refers to the utilization of natural catalysts, proteins and enzymes produced by living organisms, to perform chemical transformations on compounds of various types (Roberts *et al.*, 1995). Natural catalysts can be used in isolated form, as part of natural or engineered metabolic pathways in designated cellular hosts, or in combination with classical chemical synthesis (Coward-Kelly and Chen, 2007; Turner and Truppo, 2013). The success of biocatalysis has been accelerated by significant advances in recombinant DNA technology and synthetic biology, which enabled re-design and reconstruction of biological systems (Becker and Wittmann, 2015; Wen *et al.*, 2013). This allowed the production of several different chemicals by fermentation, in microbial hosts, by utilising sugars, proteins, lipids and organic acids, as starting material. Such processes provide a more sustainable way of manufacturing, because they do not rely on oil derivatives but rather on inexpensive and renewable carbon feedstocks (Liese *et al.*,

2007; Roberts *et al.*, 1995). Although limitations are still present, such as the ability to produce stable enzymes or substrate/product and diffusion limitations in whole-cell systems (Coward-Kelly and Chen, 2007; Reetz, 2013), achievements in metabolic and protein engineering have contributed to partially cross these natural boundaries (Bornscheuer *et al.*, 2012; Davids *et al.*, 2013; Marrs *et al.*, 1999; Otte and Hauer, 2015), and the last decades have seen a substantial increase in the application of biocatalysis (Becker and Wittmann, 2015; Liese *et al.*, 2007).

A process for the complete bioproduction of MMA and MAA has not been reported to date and no real data that can be compared to the chemical counterparts is available. Chemoenzymatic processes have been attempted, where intermediates, 2-methyl-1,3-propanediol (Pyo *et al.*, 2012) or 2-butanone (Eastham *et al.*, 2013), are first produced by biocatalysis and then converted to MAA chemically. Although these examples rely on non-petrochemical starting material, they still require high temperature, energy intensive chemical reactions which may have a significant impact on costs.. Moreover, efficiencies are poor. Attempts to directly produce MAA by fermentation of renewable sources in engineered microbial hosts have also been reported, but production yields are not disclosed (Sato *et al.*, 2014a; Sato *et al.*, 2014b). A significant complication is that, in a fermentative or biocatalytic framework, the production occurs in water at nearly ambient temperature. Formation of MMA in such an environment is disfavoured, as the reaction thermodynamically favors hydrolysis of the ester ($\Delta G = -5$ kcal mol^{-1}). To overcome this, conventional water-based esterification relies on high temperatures and acidic conditions (Nelson *et al.*, 2008; Sakamuri, 2000), but this is incompatible with biocatalysis. Therefore, MAA represents a better target than MMA for the development of a bioproduction process.

2.1.3 Potential routes

Although some MAA ester derivatives have been found in oil extracts of roman chamomile (Antonelli and Fabbri, 1998), MAA does not occur in nature. Therefore, the only possible route to produce it biologically is to find an enzyme that is able to catalyse its formation from other naturally occurring metabolic intermediates. Preferably, this conversion should occur in one step and, ideally, it should be possible to enhance the formation of the intermediate by synthetic biology. Examples of these intermediates are numerous (Figure 2.5). 3-Hydroxybutyrate and methacrylyl-CoA, both part of the valine degradation pathway (Marshall and Sokatch, 1972; Massey *et al.*, 1976; Robinson *et al.*, 1957), could potentially be converted to MAA by dehydration or hydrolysis, and enzymes of the class hydro-lyases (EC 4.2.1) or thiolester hydrolases (EC 3.1.2) could potentially catalyse the reaction. 3-Aminoisobutyrate, also part of the valine degradation pathway and found in the reductive thymine degradation route as well (Vogels and Van der Drift, 1976), could potentially form MAA by deamination/dehydration, and ammonia-lyases (EC 4.3.1) could potentially function as catalysts.

Alternatives are the decarboxylation of itaconic, mesaconic, citraconic or citramalic acid (Figure 2.5). Itaconic, mesaconic and citraconic acid could potentially be converted to MAA in a single-step decarboxylation, whereas citramalic acid requires both a dehydration and a decarboxylation step. The enzymatic class that needs to be searched is that of decarboxylases (EC 4.1.1). The production of these organic acids has been more extensively investigated (see below) compared to 3-hydroxyisobutyrate, 3-aminoisobutyrate and methacrylyl-CoA, and, therefore, this approach presents an attractive option to explore.

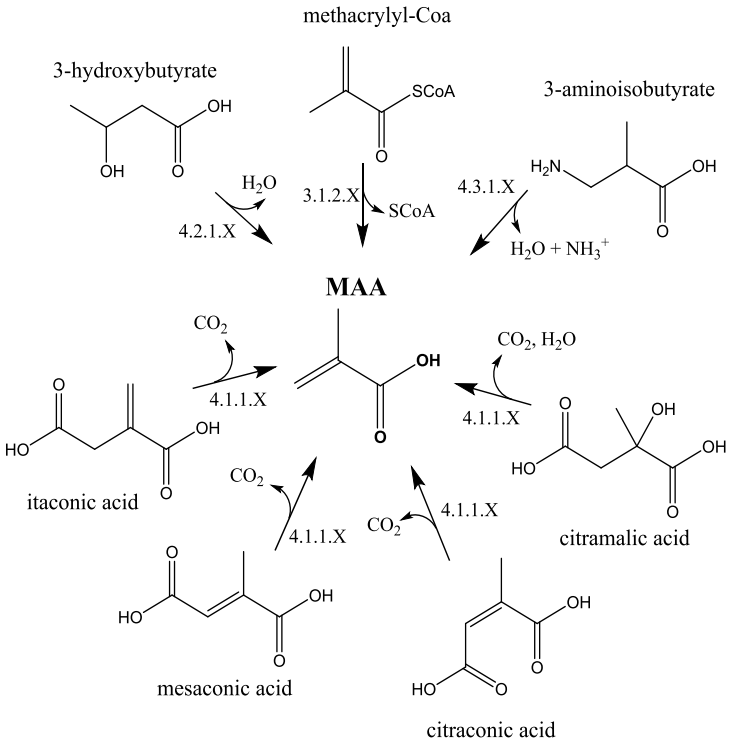


Figure 2.5: Potential biocatalytic routes to MAA
X as final EC serial number indicates that there is no known enzyme active on that specific substrate

2.1.4 Bioproduction of candidate substrates

2.1.4.1 Itaconic acid

Itaconic acid is an unsaturated dicarboxylic acid with a wide variety of applications in the production of resins, and as a building block for acrylate polymers (Okabe *et al.*, 2009; Steiger *et al.*, 2013). Historically obtained by thermal decomposition of citric acid, since the 1960s, its production is achieved by submerged fermentation of carbohydrates by the fungus, *Aspergillus terreus* (Okabe *et al.*, 2009; Steiger *et al.*, 2013). Although itaconic acid can also be naturally produced by other microorganisms such as *Ustilago sp., Candida sp.* and *Rhodotorula sp.* (Willke and Vorlop, 2001), *A. terreus* has remained the predominant production host and several attempts to enhance productivities and production titres have been reported. This included metabolic flux modifications (Tevž *et al.*, 2010) and resolving product toxicity issues (Kobayashi and Nakamura, 1964). To date, the maximum production titres reached

are around 80 g/L in 6 days using *A. terreus* (Okabe *et al.*, 2009; Steiger *et al.*, 2013; Yahiro *et al.*, 1997).

The main enzyme involved in itaconic production is *cis*-aconitate decarboxylase (CAD) (Bentley and Thiessen, 1957; Dwiarti *et al.*, 2002; Kanamasa *et al.*, 2008; Li *et al.*, 2011). CAD converts the TCA cycle intermediate, *cis*-aconitic acid, to itaconic acid in the mitochondria, and itaconic acid is then released in the cytosol and outside the cells (Figure 2.6). Attempts have been made to engineer other heterologous hosts with CAD. Engineered *A. niger* was able to successfully produce up to 7 g/L itaconic acid in 78 h, when putative mitochondrial transporters were also expressed and enzyme production was specifically targeted to the mitochondria (Blumhoff *et al.*, 2013; Li *et al.*, 2013a; Li *et al.*, 2013b; van der Straat *et al.*, 2014). Examples of engineered *Saccharomyces cerevisiae* and *E. coli* were also reported (Blazeck *et al.*, 2014; Hiller *et al.*, 2014; James and Pei-Ching, 2010; Okamoto *et al.*, 2014; Vuoristo *et al.*, 2015). With optimisation of metabolic flux, production reached levels up to about 4.34 g/L in 105 h using *E. coli* (James and Pei-Ching, 2010; Okamoto *et al.*, 2014). More recently, itaconic acid fermentation has also been demonstrated in *Corynebacterium glutamicum*, with a final production of 7.8 g/L in 72 h (Otten *et al.*, 2015). Although the final production titres are lower than those achieved with *A. terreus* (80 g/L), these heterologous hosts can produce itaconic acid quicker. Therefore, with more development, there is room for further improvement to the already efficient current biproduction process using *A. terreus*.

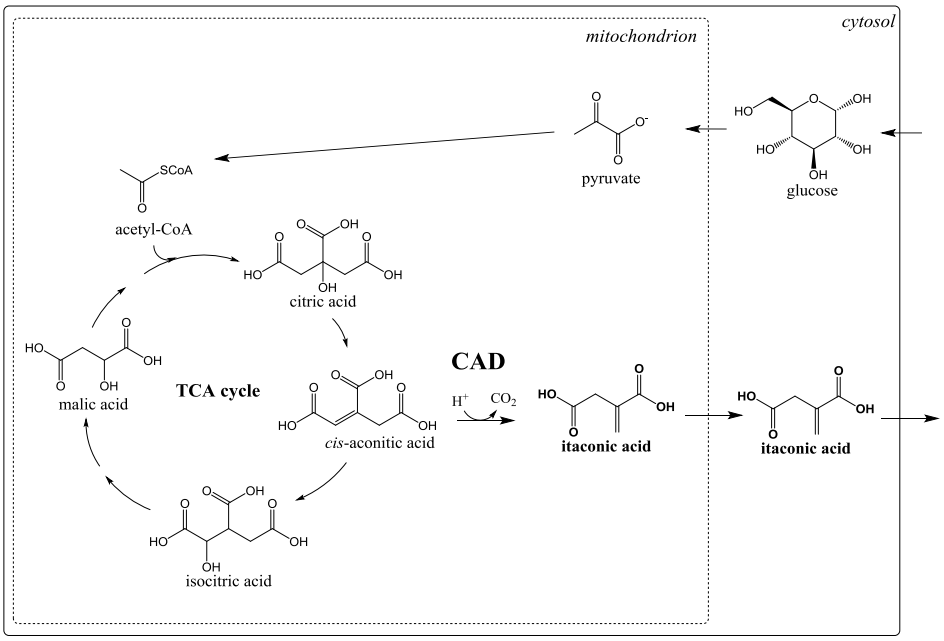


Figure 2.6: Itaconic acid formation in *A. terreus*
In the cytosol, *via* glycolysis, glucose is converted to pyruvate, which subsequently enters into the mitochondrion. Pyruvate is converted to acetyl-CoA which enters the TCA cycle. *cis*-Aconitic acid, an intermediate of the TCA cycle, is then converted to itaconic acid by *cis*-aconitate decarboxylase (CAD). Itaconic acid exits the mitochondrion and is further released outside the cell.

2.1.4.2 Mesaconic acid

Mesaconic acid is another isomeric unsaturated dicarboxylic. It is used as a fire retardant or as a precursor to other compounds. Although its production still relies on chemical synthesis from citric acid, it has also been isolated from the fermentative bacterium *Clostridium tetanomorphum*, where it is an intermediate of the glutamate degradation pathway (Figure 2.7) (Wachsman, 1956). In this pathway, L-glutamate is converted to pyruvate and acetate in four steps, and pyruvate is subsequently used in the TCA cycle. 3-Methlyaspartate is first produced from glutamate by glutamate mutase (GM). Mesaconic acid is then formed through deamination of 3-methylaspartate (Goda *et al.*, 1992; Kato and Asano, 1997), catalysed by 3-methyl aspartase (3-MA). This pathway is present in other organisms such as *Citrobacter sp., Klebsiella planticola* and *Morganella morganii* (Goda *et al.*, 1992; Kato and Asano, 1997). Although normally an intermediate of the pathway, mesaconic acid can accumulate in a mixture with other by-products, such as acetate and butyrate (Wachsman, 1956). First attempts to produce

mesaconic acid using *E. coli* engineered with 3-MA and GM has also been reported, with production titres of up to around 8 g/L (Eastham *et al.*, 2015; Wang and Zhang, 2015).

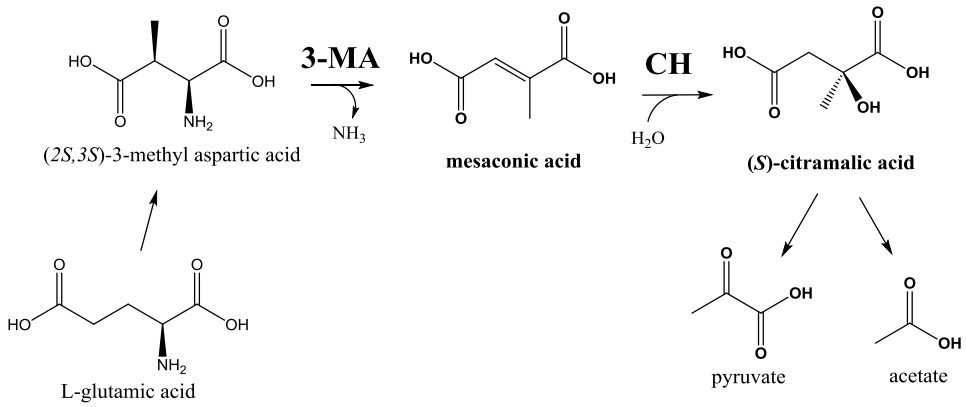


Figure 2.7: Glutamate degradation pathway
3-Methyl aspartase (3-MA) catalyses formation of mesaconic acid and citramalate hydrolase (CH) catalyses the formation of (*S*)-citramalic acid.

Mesaconic acid is also an intermediate of the citramalate cycle in *Rhodospirillum rubrum* (Berg and Ivanovsky, 2009), an anaplerotic route to convert acetyl-CoA and pyruvate to glyoxylate and propionyl-CoA (Figure 2.8). This pathway has been less extensively studied than the glutamate degradation pathway, and not all of the enzymes have been characterised. More recently, the biocatalytic production of mesaconic acid from mesaconyl-CoA has been reported (Sonntag *et al.*, 2014). Mesaconyl-CoA is a natural intermediate of the ethylmalonyl-CoA pathway in *Methylobacterium extorquens*, and the introduction of exogenous thioesterases led to the production of mesaconic acid *in vivo*.

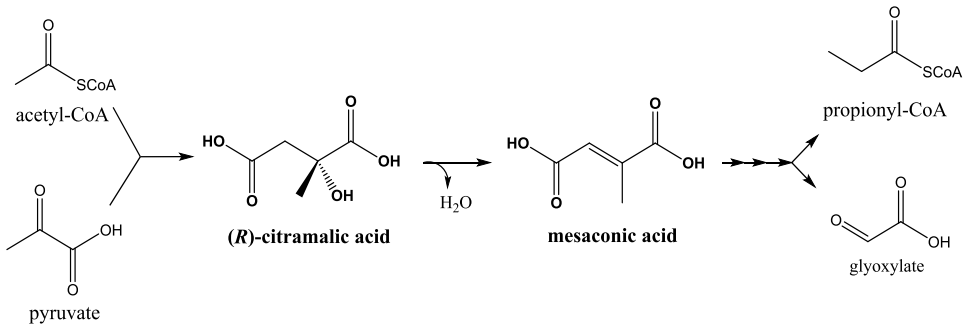


Figure 2.8: Citramalate cycle in *R.rubrum*

2.1.4.3 Citramalic acid

Citramalic acid is a chiral analogue of malic acid which is mostly used for the preparation of chiral aliphatic sulfones (Kutner *et al.*, 1996) and isocyanates (Pires and Burger, 1996). Different routes have been identified for its biocatalytic production. The (*R*)- and (*S*)-enantiomers are present as metabolic intermediates of several naturally occurring pathways. (*S*)-Citramalic acid is found in the glutamate degradation pathway (Figure 2.7) (Wachsman, 1956), where it is formed by hydration of mesaconic acid catalysed by citramalate hydrolase (CH) and it is further converted to pyruvate and acetate. (*S*)-Citramalic acid is also found in the itaconic acid degradation pathway discovered in *Alcaligenes sp.* and *Pseudomonas sp.* (Figure 2.9) (Bing fang *et al.*, 2000; Cooper and Kornberg, 1964; Ueda *et al.*, 1993).

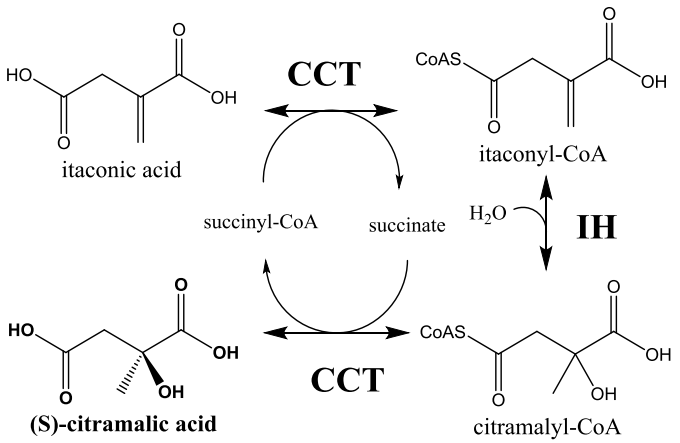


Figure 2.9: Itaconic degradation pathway in *Alcaligenes sp*. and *Pseudomonas sp*. Citramalate-CoA transferase (CCT) catalyses the conversion of itaconic acid to itaconyl-CoA, and the conversion of citramalyl-CoA to (*S*)-citramalic acid. Itaconyl-CoA hydratase (IH) catalyses the conversion of itaconyl-CoA to citramalyl-CoA .

The enzymes involved in this pathway have not yet been characterised, but it is believed that itaconic acid is converted to citramalic acid through the formation of their CoA derivatives, itaconyl-CoA and citramalyl-CoA. In this hypothesis, the enzymes involved are citramalate-CoA transferase (CCT) and itaconyl-CoA hydratase (IH). Resting cells of *Al. xylosoxydans* IL142 were able to produce up to 65 g/L citramalic acid with a conversion yield of >90% (Bing fang *et al.*, 2000). Production of (*S*)-citramalic acid has

also been reported with engineered *E. coli*, but production titres were lower (Eastham *et al.*, 2015).

(*R*)-Citramalic acid is found in the *R. rubrum* citramalate pathways (Figure 2.8) (Berg and Ivanovsky, 2009) and in the isoleucine biosynthesis pathway of *Leptospira interrogans* and *Methanococcus jannaschii* (Figure 2.10) (Howell *et al.*, 1999; Xu *et al.*, 2004). In both pathways, (*R*)-citramalic acid is produced by condensation of acetyl-CoA and pyruvate, catalysed by citramalate synthase (CS). Enhancements *via* directed evolution prevented feedback inhibition (Atsumi and Liao, 2008) and, recently, our group has developed a process for production of (*R*)-citramalic acid from glucose using *E. coli* engineered with the modified CS (Eastham *et al.*, 2015). The final production titre achieved was over 100 g/L.

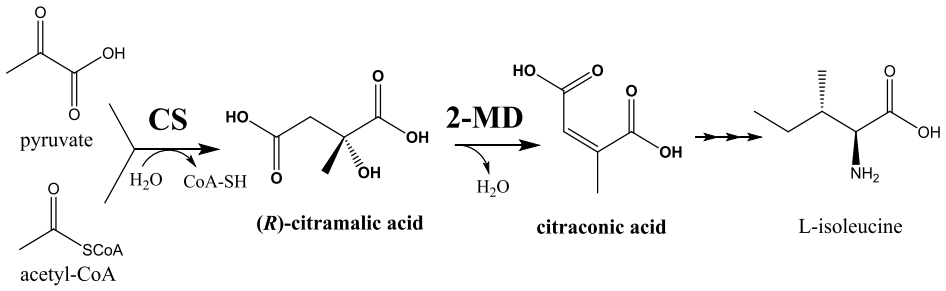


Figure 2.10: Isoleucine biosynthesis pathway
Citramalate synthase (CS) catalyses the formation of (*R*)-citramalic acid from pyruvate and acetyl-CoA. (*R*)-2-methylmalate dehydrogenase (2-MD) converts (*R*)-citramalic acid to citraconic acid.

2.1.4.4 Citraconic acid

Citraconic acid is the isomer of citric acid, and is produced chemically from citric acid (Williams, 2006). There are no reports of its production through fermentation or biotransformation. Nevertheless, as with citramalic acid, citraconic acid is an intermediate of the isoleucine biosynthesis pathway (Xu *et al.*, 2004) (Howell *et al.*, 1999), where it is produced from (*R*)-citramalic acid by (*R*)-2-methylmalate dehydratase (2-MD) (Figure 2.10). Since the fermentative production of (*R*)-citramalic acid has already been reported (see above), production of citraconic acid could be

obtained in principle by adding the step catalysed by 2-methylmalate dehydrogenase.

The technology to produce metabolic intermediates that can be converted to MAA is available. The biological production of itaconic, mesaconic, citramalic and citraconic acid has been reported, and some of them can be produced by fermentation at high concentrations. Therefore, the formation of MAA through enzymatic decarboxylation of these substrates is very attractive and presents, in principle, a straightforward route to develop the first MAA bio-based production process. Nevertheless, there is still the need to find an enzyme that can catalyse the final, crucial step. If this enzyme can be found, synthetic biology tools can be used to design a whole metabolic pathway for the direct bioproduction of this important commodity chemical.

2.1.5 Use of decarboxylases

In order to convert itaconic, mesaconic, citraconic or citramalic acid to MAA, an enzyme is required to catalyse the removal of a CO_2 group. In nature, this reaction is catalysed by enzymes that are categorised under the class of decarboxylases (EC 4.1.1). Decarboxylases are involved in the metabolism of essentially all nutrients that serve as sources of energy in living organisms (Li *et al.*, 2012b) and, to date, almost 100 different sub-classes have been identified (IUBMB). Most decarboxylases utilise a cofactor. The cofactor can be either organic (biotin, pyridoxal-5-phosphate, pyruvoyl, thiamine diphosphate, flavin and NAD/NADP), inorganic (iron, zinc, manganese *etc*.) or both. Decarboxylases which do not require any cofactor have also been discovered (Cuia *et al.*, 1999; Jordan and Patel, 2013).

Biotin-dependent decarboxylases (*e.g.* glutaconyl-CoA decarboxylases, malonate decarboxylase and oxaloacetate decarboxylase) typically catalyse the decarboxylation of α-keto acids or thioesters coupled with sodium transport from the cytoplasm into

the periplasm of some bacteria and archaea, and biotin serves as a transient carboxyl carrier (Buckel, 2001) (Figure 2.11). Sometimes, divalent metals are needed to correctly position the substrate.

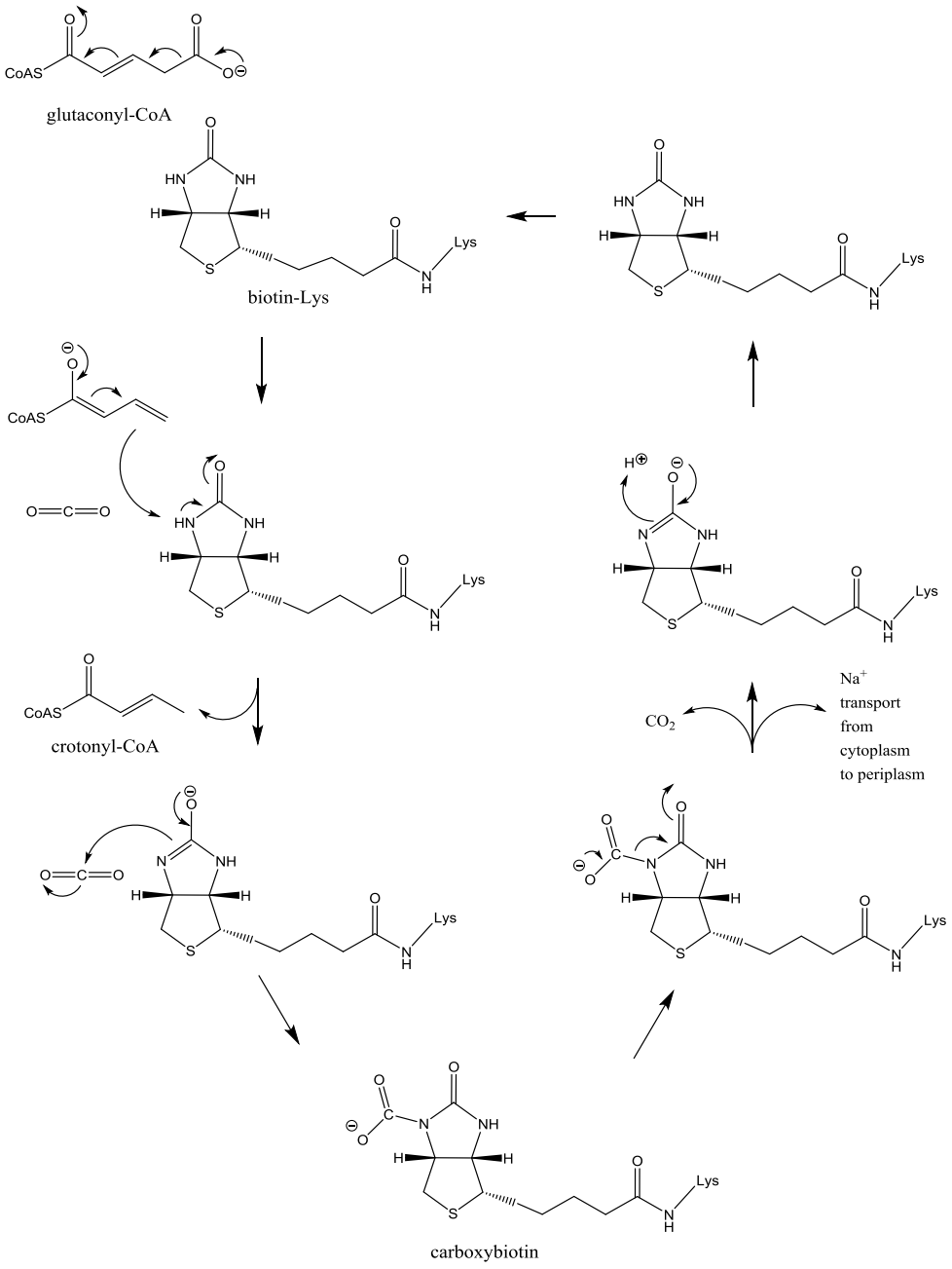


Figure 2.11: Biotin-dependent decarboxylation
The example shown is that of glutaconyl-CoA decarboxylase. Biotin is linked to a lysine residue (biotin-Lys). The primary substrate is initially polarised in the active site of the enzyme and induces transient decarboxylation. The acidic N1 of biotin promotes the proton transfer to the anion of the substrate and crotonyl-CoA is released. CO_2 is then fixed to biotin. In the last steps of the reaction, carboxybiotin is decarboxylated coupling the transfer of sodium ions from the cytoplasm to the periplasm, driven by the protonation of the 2-carbonyl oxygen of carboxybiotin. Biotin finally returns to its initial form (Wendt *et al.*, 2003).

Pyridoxal-5-phosphate (PLP)-dependent decarboxylases (*e.g.* glutamate decarboxylase, aromatic L-amino acid decarboxylases *etc.*) can catalyse a wide range of decarboxylations, including amine and amino acids (Alexander *et al.*, 1994). PLP, in the form of an aldimine with a lysine residue, serves as an "electron sink" to stabilize the carbanion formed as the transition state before the loss of CO_2 (Toney, 2005) (Figure 2.12). Very similar behaviour is that of pyruvoyl-dependent decarboxylases (*e.g.* arginine decarboxylase, histidine decarboxylase, aspartate decarboxylase *etc.*), where the pyruvoyl group covalently bound to a serine functions as the carbanion stabilizer (Webb and Abell, 2003).

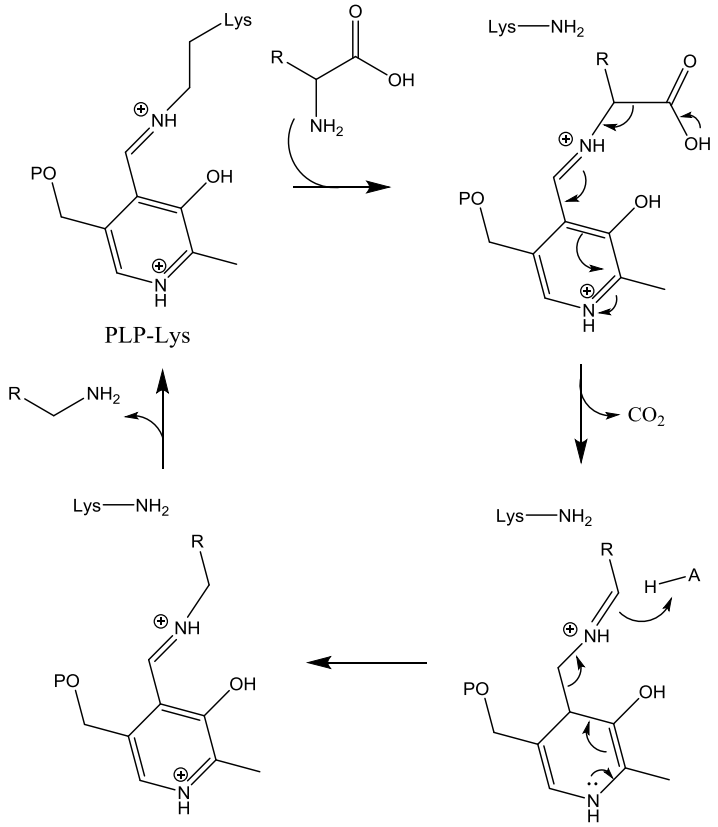


Figure 2.12: PLP-dependent decarboxylation
Pyridoxalphosphate (PLP) is linked in the form of an aldimine with the ε-amino group of a lysine (Lys) residue. PLP-Lys is subsequently transaldiminated to the substrate. This stabilises the carbanion created during the transition state intermediate which leads to decarboxylation. A general acid facilitates the protonation at Cα to form an imine. The imine finally undergoes transaldimination with Lys and the amine product is released.

Thiamine diphosphate (ThDP)-dependent decarboxylases (*e.g.* pyruvate decarboxylase, benzoylformate decarboxylase,

phenylpyruvate decarboxylase *etc.*) typically catalyse the decarboxylation of α-keto acids (Jordan, 2004; Versées *et al.*, 2007). They usually contain a divalent cation as cofactor as well, which serves to coordinate the diphosphate moiety of ThDP (Jordan, 2004; Versées *et al.*, 2007). ThDP binds the carbonyl group of the keto acid and stabilises the negative charge generated by the removal of CO_2 (Figure 2.13).

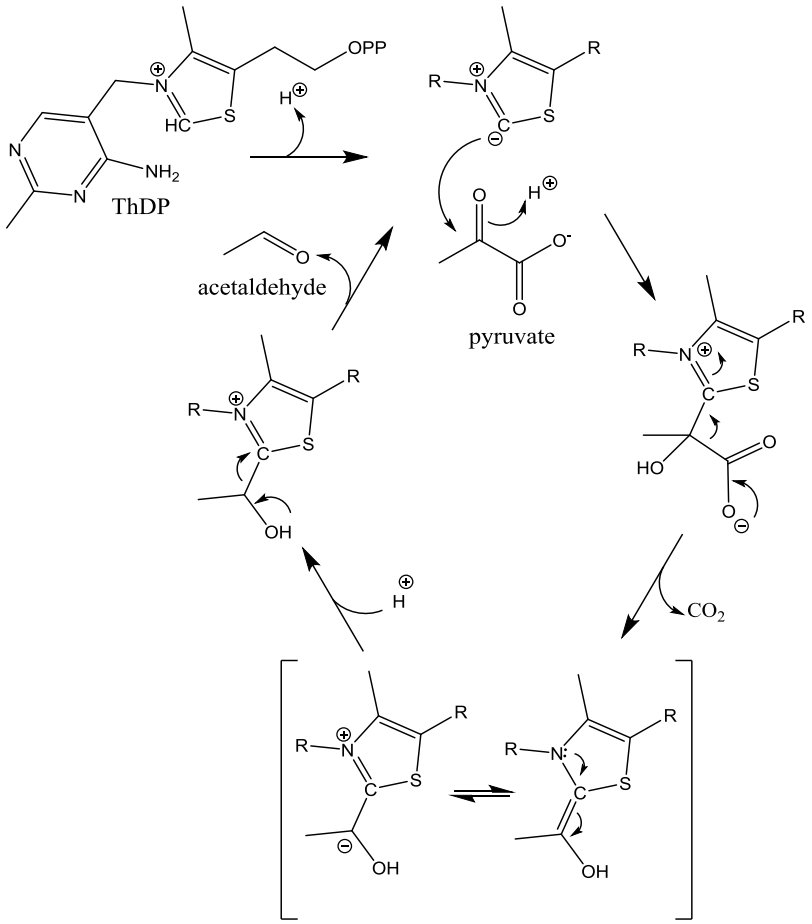


Figure 2.13: ThDP-dependent decarboxylation
The example shown is that of pyruvate decarboxylase. The thiazolium C2 atom of thiamine diphosphate (ThDP) is deprotonated to an ylide. The ylide is very nucleophilic and attacks the carbonyl of the primary substrate to form a covalent tetrahedral intermediate. This intermediate undergoes decarboxylation and a carbanion, stabilised by resonance with its enamine, is formed. The thiazolium ring of ThDP acts as an electron sink. Protonation on the carbon of the enamine leads to a second tetrahedral intermediate which is subsequently deprotonated with concomitant release of the product, acetaldehyde (Lie *et al.*, 2005).

Flavin-dependent decarboxylases include different families of cysteine decarboxylases (homo-oligomeric flavin-containing cysteine decarboxylases, phosphopantothenoylcysteine decarboxylases and *L*-cysteine decarboxylase) (Hernández-Acosta *et al.*, 2002; Strauss

et al., 2004). The flavin cofactor, FAD or FMN, gets reduced with concomitant oxidation of the cysteine moiety, which favours the decarboxylation (Figure 2.14) (Strauss *et al.*, 2004). The electron transfer mechanism has not yet been completely understood.

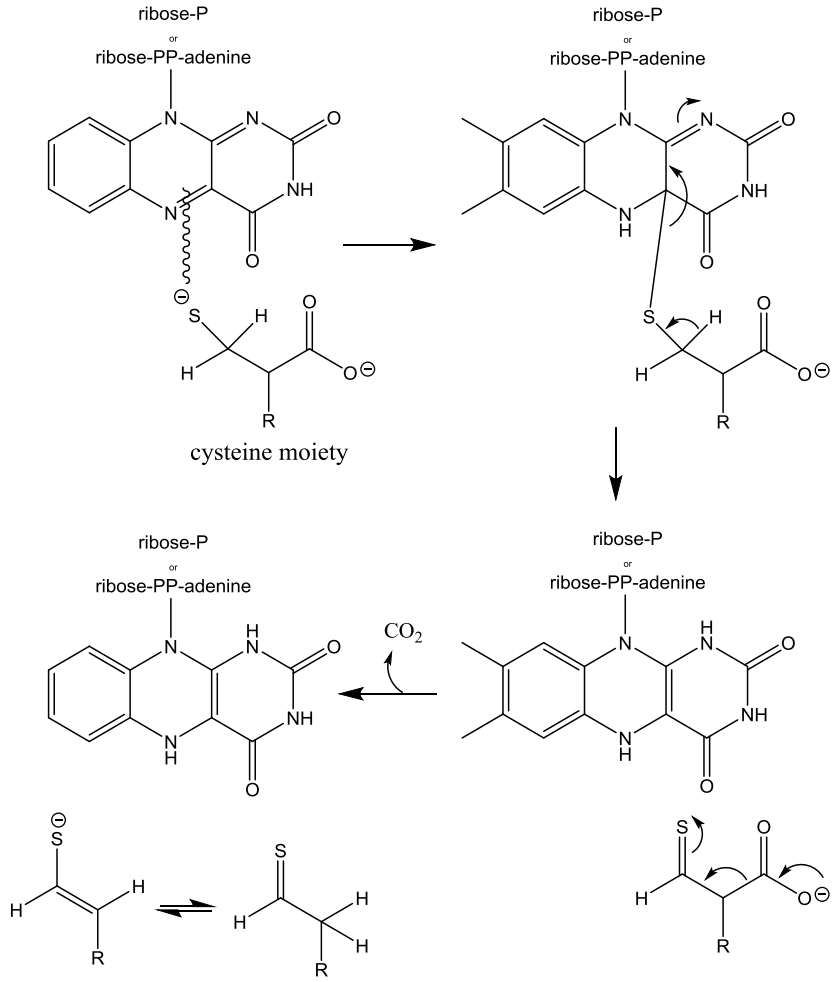


Figure 2.14: Flavin-dependent decarboxylation
In the proposed model, the thiol side chain of the cysteine moiety is oxidized to a thioaldehyde with the concomitant reduction of the flavin cofactor, FAD or FMN. Decarboxylation occurs spontaneously favoured by the delocalization of the negative charge and the enethiol is released as final product.

Even for NAD/NADP-dependent decarboxylases (*e.g.* malic enzymes, tartrate dehydrogenase, isocitrate dehydrogenase *etc.*), the decarboxylation occurs in an oxidative fashion but the mechanism is still unclear (Grissom and Cleland, 1988; Tao *et al.*, 2003). It is believed that NAD, or NADP, serves as acceptor for the proton transfer from the primary substrate, which subsequently drives the decarboxylation (Figure 2.15). Many of these

decarboxylases require a divalent metal ion as additional cofactor, others do not require any instead (Bar-Peled *et al.*, 2001).

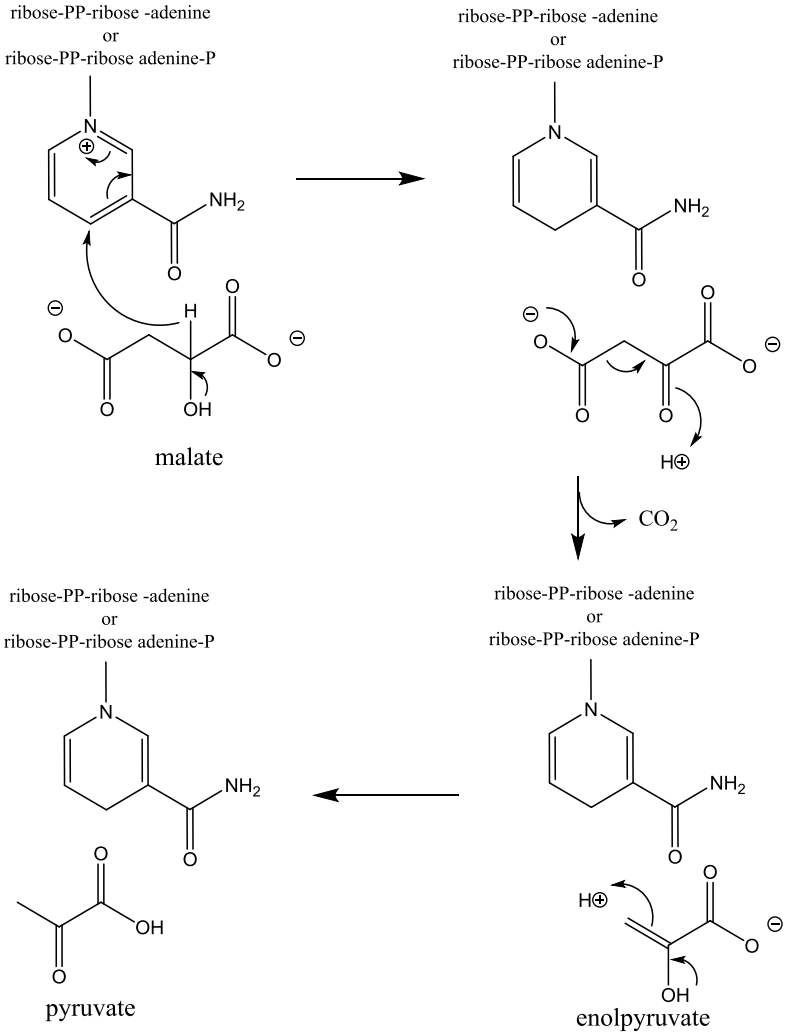


Figure 2.15: NAD/NADP-dependent decarboxylation
The example shown is that of malic enzyme. The β-hydroxyl group is deprotonated and this facilitates a hydride transfer from the primary substrate to the NAD (or NADP), leading to dehydrogenation. Subsequent decarboxylation forms enolpyruvate and then, after another proton transfer to the resulting enolate group, pyruvate (Liu *et al.*, 2000).

Oxalate decarboxylase and α-amino-β-carboxymuconate-ε-semialdehyde decarboxylase (ACMSD) are two examples of decarboxylases that require inorganic cofactors as sole cofactors. Oxalate decarboxylase catalyses the conversion of oxalate to formate and CO_2, utilising Mn^{2+} and O_2 (Tanner *et al.*, 2001) (Figure 2.16). ACMSD requires Zn^{2+} as activator of a water molecule that subsequently facilitate the nucleophilic attack on the primary substrate (Liu and Zhang, 2006) (the mechanism is extensively

explained later in the text). Finally, two examples amongst the few decarboxylases which function without any exogenous cofactors are orotidine monophosphate decarboxylase (Cuia *et al.*, 1999) and arylmalonate decarboxylase (Miyamoto and Ohta, 1992). These enzymes catalyse non-oxidative decarboxylations but, although several insights have been reported (Fujihashi *et al.*, 2013; Goryanova *et al.*, 2013; Jamshidi *et al.*, 2015; Jamshidi *et al.*, 2014; Lee, 1997; Lewin *et al.*, 2015; Lind and Himo, 2014), the reaction mechanism remains unclear.

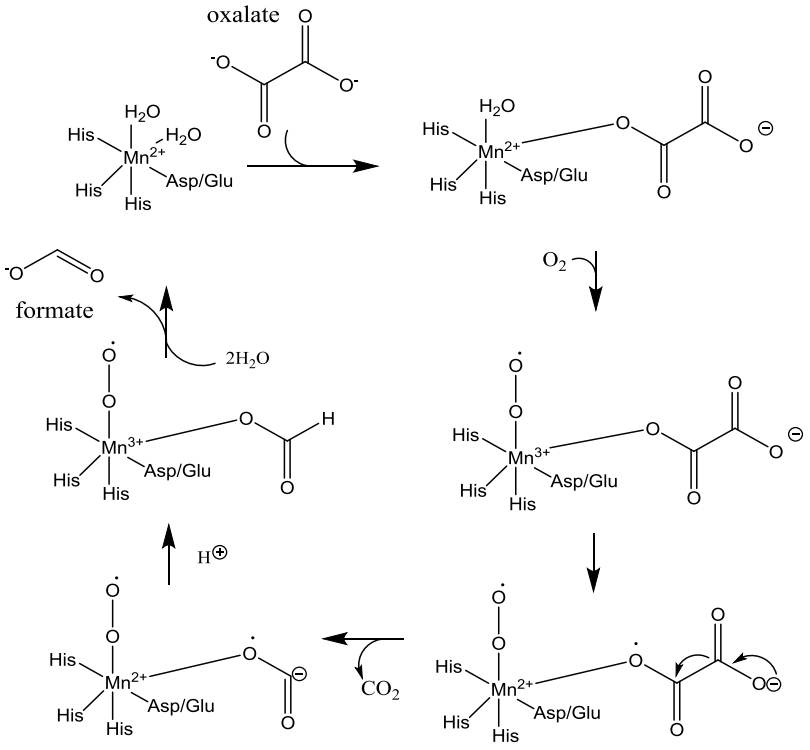


Figure 2.16: Oxalate decarboxylation
Mn^{2+} is coordinated in the active site by three histidine (His) residues, one acidic residue such as aspartate (Asp) or glutamate (Glu), and two H_2O. Oxalate binds to the metal ion replacing one of the water molecule ligands. Then O_2 binds to Mn^{2+} replacing the second water molecule and forming Mn^{3+}. An electron is transferred from the substrate to the metal and a radical is generated. CO_2 is released and the intermediate, bound to the metal, present a carbanion. The latter is protonated in concomitance to an electron transfer from Mn^{2+}. The product, formate, is finally released and the metal regenerated.

Decarboxylases provide a significant variation of cofactor dependence and reaction chemistries (Jordan and Patel, 2013; Li *et al.*, 2012b) which enables them to act on a wide range of substrates. For this reason, they can be a very useful tool for the defunctionalisation of organic molecules, and have been employed in

biocatalysis for the synthesis of alcohols, carboxylic acids, diamines, olefins and fine chemicals (Kourist *et al.*, 2014). Although there is no class of decarboxylases that specifically act on mesaconic, itaconic, citraconic or citramalic acid, it is possible, in principle, to exploit promiscuous enzymes with reaction chemistries similar to those required for the target reactions.

Nevertheless, amongst the main sub-classes of decarboxylases, there is no one that raises particular interest. Biotin-dependent, ThDP-dependent and oxalate decarboxylases act on keto acids, and the carbonyl of the substrate usually plays a crucial role in the reaction (Figure 2.11, Figure 2.13, Figure 2.16). Mesaconic, itaconic and citraconic acid are not keto acids. PLP/pyruvoyl-dependent and flavin-dependent decarboxylases require a substrate with an amine and a thiol, respectively (Figure 2.12, Figure 2.14). None of the substrates of interest presents such moieties. NAD/NADP-dependent decarboxylases, although they act on dicarboxylic hydroxy acids (Figure 2.15), similar in structure to citramalic acid, only perform decarboxylation. Both decarboxylation and dehydration are required for conversion of citramalic acid to MAA. Therefore, candidate enzymes for the decarboxylation of itaconic, mesaconic, citraconic and citramalic acid must be searched within other sub-classes. Although these sub-classes are often less characterised, they could potentially provide better alternatives and a broader substrate range, and this requires further investigation.

2.2 Protein engineering

Activity of an enzyme on non-natural substrates is typically significantly lower than that on its natural substrate. Therefore, if a decarboxylase with substrate promiscuity able to act on itaconic, mesaconic, citraconic or citramalic acid can be found, it is very likely that its efficiency will need improving. It is also possible that a suitable decarboxylase cannot be found and, in that case, it will be necessary to develop a new enzyme that is able to catalyse a non-

natural reaction. Protein engineering provides promising tools in both cases (Bornscheuer and Kazlauskas, 2004).

Protein engineering refers to the procedures by which proteins and enzymes are modified to result in new or improved functions (Ulmer, 1983). This is achieved by introducing diversity in the amino acid sequence, which can lead to new, useful structural, binding, stability and catalytic properties. Protein engineering contributed to develop several enzyme-based processes that are industrially viable, and there has been a tremendous increase in novel concepts and techniques (Bornscheuer *et al.*, 2012; Davids *et al.*, 2013; Ruff *et al.*, 2013; Toogood and Scrutton, 2013). "In the past, an enzyme-based process was designed around the limitations of the enzyme; today, the enzyme is engineered to fit the process specifications" (Bornscheuer *et al.*, 2012). Successful examples of protein engineering on decarboxylases to improve activity or to change substrate specificity have been reported (Gogerty and Bobik, 2010; Miyamoto, 2014; Miyauchi *et al.*, 2011; Morley *et al.*, 2013; Okrasa *et al.*, 2009).

2.2.1 Screening and selection

In protein engineering, everything starts from the original gene sequence encoding the target enzyme. DNA variations are introduced in the gene to alter the amino acid sequence when the enzyme is expressed. This can include simple mutations, insertions or deletions of amino acids. The gene products are then selected or screened for improved target properties (Davids *et al.*, 2013), and more and more variations can potentially be introduced for further enhancements. The screening and selection of improved enzyme variants is a bottle neck to be successful. If an assay cannot be developed, protein engineering cannot be performed. The number of variants screened is also very important. High throughput assays enable screening or selection of a large number of variants and therefore, evaluation of a high level of variation. This increases the likelihood of finding enzymes with improved target properties

(Goddard and Reymond, 2004). In contrast, low throughput assays restrict the number of variants analysed and, consequently, the level of variation explored.

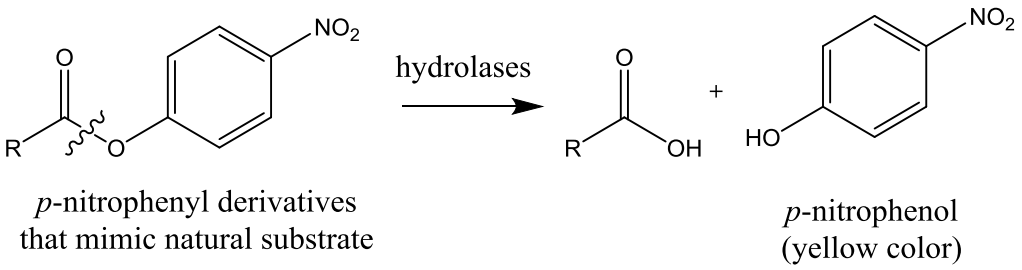


Figure 2.17: Colorimetric reaction based on *p*-nitrophenyl derivatives
Enzymatic hydrolysis of the *p*-nitrophenyl esters generates *p*-nitrophenol.

Screening assays evaluate each individual enzyme variant (Arnold and Georgiou, 2003; Xiao *et al.*, 2015). They are usually performed *in vitro*, and enzyme variants are expressed in host cells and subsequently tested in crude extracts after cell lysis. The target reaction is then monitored and analysed with colorimetric/fluorimetric assays, exploiting absorbance and fluorescence properties of substrates and products (Goddard and Reymond, 2004). An example is the utilisation of *p*-nitrophenyl derivatives to detect the activity of hydrolases. The *p*-nitrophenyl derivatives mimic the natural substrates of the enzyme but add a *p*-nitrophenyl moiety at the cleavage site. When the enzyme is active, *p*-nitrophenol is released and can be detected by its yellow color (Kurioka and Matsuda, 1976; Prasad *et al.*, 2011) (Figure 2.17). When chromogenic substrates or products are not applicable, indirect colorimetric reactions can be exploited. A general example is the utilisation of pH indicators, where the release of protons during an enzymatic reaction can be monitored through a change in the color of the indicator added to the reaction mixture (Chapman and Wong, 2002; Janes *et al.*, 1998; Rosenberg *et al.*, 1989; Yi *et al.*, 2012). The utilisation of multi-well plates and automated equipment increases the number of reactions that can be monitored simultaneously, and certain screening assays can analyse up to 10^8 enzyme variants per day (Reymond, 2005; Yang and Withers,

2009). Nevertheless, due to the chemistry and availability of suitable substrates/products, direct or indirect colorimetric/fluorimetric reactions are often not applicable. In these cases, the ability to develop a valid screening assay is limited.

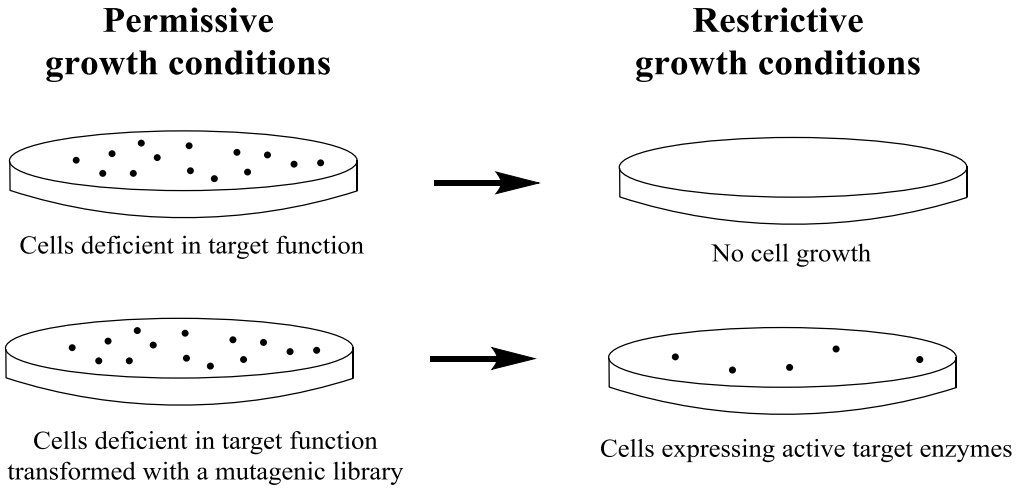


Figure 2.18: Principles of selection assays

In selection assays, evaluation of each individual enzyme variant is not required. Non-improved (or non-functional) enzyme variants are directly eliminated *in vivo* by applying a certain selective pressure to the cells expressing the enzyme (Arnold and Georgiou, 2003; Reymond, 2005). Cells lacking or deficient in a specific function, which is linked to their ability to grow at restrictive conditions, are transformed with a mutagenic library. Only cells expressing an active enzyme that complements this function will grow (Figure 2.18). A good example is the selection of an enantioselective lipase from *Bacillus subtilis* (Boersma *et al.*, 2008). An aspartate auxotroph *E. coli* was used as a host cell for a mutant library of *B. subtilis* Lipase A. The strain was then plated on minimal medium supplemented with the aspartate ester of the target enantiomer. As a result, only cells expressing a lipase that could hydrolyse this substrate were able to release aspartate and to grow at those conditions. Selection assays enable to assess much larger libraries than random screening assays, as only positive variants can grow and be selected for further rounds of protein engineering (Arnold and Georgiou, 2003). Nevertheless, it is not possible to

establish a complementation system for every target enzymatic reaction, and this represents a significant limitation.

Examples of screening for decarboxylses have been reported, but they are always limited to specific enzymes. An example is the use of fluorogenic reagents that react with the amine group of amino acids only after decarboxylation by amino acid decarboxylases (Medici *et al.*, 2011). Neverthless, a general screening or selection method for decarboxylases has not been reported to date. Therefore, in the case that protein engineering is needed to develop an enzyme that converts itaconic, mesaconic, citraconic or citramalic acid to MAA, effort must be spent in developing an assay. Without a valid and robust screening or selection assay in place, protein engineering cannot be performed and, therefore, this step is crucial for the project. Focus should be given to the detection of MAA or of CO_2, or to the monitoring of substrate consumption.

2.2.2 Introduction of variation

Another key step in protein engineering is the introduction of variation into the amino acid sequence of the target enzyme. Depending on the strategy used, different levels of variation can be generated (Arnold and Georgiou, 2003). Three main approaches can be identified: rational design, directed evolution, and semi-rational design (Arnold and Georgiou, 2003).

2.2.2.1 Rational design

In rational design, specific modifications of the amino acid sequence are introduced rationally at those positions that are believed to contribute to specific functions (Brannigan and Wilkinson, 2002; Cedrone *et al.*, 2000). Since early days, this approach has shown success (Bonagura *et al.*, 1999; Mata *et al.*, 1999; Russell and Fersht, 1987) but an extensive knowledge of the chemistry of the target enzyme is required to make decision on mutagenesis. This knowledge can be obtained with functional studies, crystal structures and homology models, or by employing

multiple sequence alignments (Folkertsma *et al.*, 2004; Kuipers *et al.*, 2010) and prediction algorithms (Chen *et al.*, 2009; Feng *et al.*, 2012). The screening effort required for rational design is minimal, since only a few enzyme variants need to be generated and can usually be tested individually without the need of high throughput screening or selection assays. Nevertheless, the required information is not available for all enzymes and, despite the continuous development of bioinformatic tools (Brooksbank *et al.*, 2014; Gasteiger *et al.*, 2003), data can often be hard to interpret.

2.2.2.2 Directed evolution

Directed evolution (Arnold, 1996) mimics the mechanisms of the natural Darwinian evolution, for which enzymes accumulate random mutations that enable them to adapt to new environments, conditions, functions and substrates (Jensen, 1976; Kuchner and Arnold, 1997). This process is reproduced *in vitro*. The evolution is forced by multiple generations of random mutagenesis or gene recombination, coupled with screening and selection for enhanced enzymes (Arnold, 1996). The power of directed evolution resides in the possibility to quickly introduce a large number of variations and to screen or select only for those that are useful.

Random mutagenesis is most commonly achieved by error prone PCR (epPCR)-based methods (McCullum *et al.*, 2010) (Figure 2.19A). Point mutations are introduced along the whole length of a target gene through a PCR process where the fidelity of the DNA polymerase is altered (McCullum *et al.*, 2010; Rasila *et al.*, 2009). The alteration is obtained by using unbalanced concentrations of reaction components or by adding disrupting agents (McCullum *et al.*, 2010). This technique is robust and simple, and successful examples have been reported (Chen and Arnold, 1993; Lingen *et al.*, 2003). However, it presents some limitations. The DNA polymerase never introduces point mutations in close proximity, making it impossible to develop adjacent amino acid substitutions (McCullum *et al.*, 2010). epPCR may randomly generate stop codons

or mutations that kill the enzymatic activity, compromising the quality of the library (Ruff *et al.*, 2013). Moreover, due to the degeneracy of the genetic code, epPCR may introduce redundant codons that affect the final diversity of the library (Eggert *et al.*, 2004).

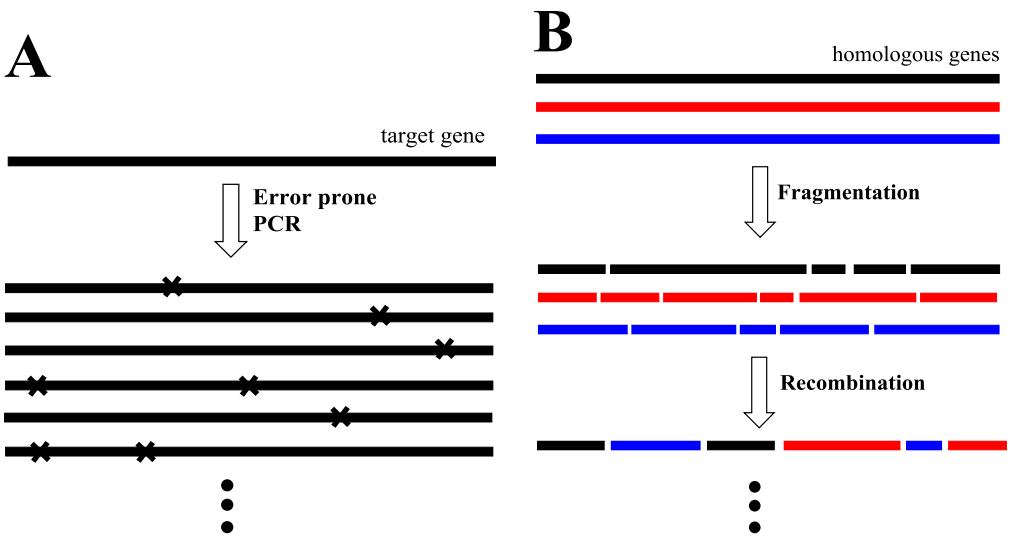


Figure 2.19: Error prone PCR and gene shuffling
Error prone PCR (A) and gene shuffling (B) methods. Mutations are indicated with an X. Different colours refer to genes from different sources.

Gene shuffling partially addresses these limitations. In this approach, diversity is generated by creating chimeric constructs through recombination of fragments, obtained from homologous genes, in a reassembly step (Stemmer, 1994) (Figure 2.19B). Homologous genes are selected from different organisms or isolated from previous engineering efforts. The obtained hybrid mutant libraries are usually highly functional because the variation derives from the recombination of functional genes and not by the insertion of potentially destabilising or redundant mutations (Coco *et al.*, 2001). Successful examples of gene shuffling include both bacterial and mammalian proteins (Crameri *et al.*, 1996a; Crameri *et al.*, 1996b; Stemmer, 1994). However, as for random mutagenesis, some limitations restrict the utilisation of this technique (Ruff *et al.*, 2013), where the most significant is the fact that recombination of

genes with low sequence identity is not efficient (Cole and Gaucher, 2011).

Directed evolution has been very successful (Bornscheuer *et al.*, 2012; Turner, 2009). However, it generates a large number of variants and often requires screening for up to 10^{10} clones (Arnold and Georgiou, 2003). This necessitates ultra-high throughput screening assays or selection procedures. When this is not possible for the target reaction, directed evolution becomes impractical (Goddard and Reymond, 2004; Reymond, 2005).

2.2.2.3 Semi-rational design

Semi-rational design (Chica *et al.*, 2005), often referred to as focused mutagenesis, combines the advantages of rational design and directed evolution (Lutz, 2010). A limited number of residues are initially selected on the basis of their possibility to affect the target properties. Random mutagenesis is then performed only at those positions (Davids *et al.*, 2013; Ruff *et al.*, 2013). This approach minimises the possibility to introduce disruptive mutations, and narrows down screening efforts by reducing the number of mutations introduced (Reetz *et al.*, 2008). The number of variants required can go down to 10^3 or 10^2 (Reetz *et al.*, 2008). Successful examples of semi-rational design include enhanced enzyme activity (Chica *et al.*, 2005; Conti *et al.*, 2014), selectivity (Nobili *et al.*, 2013) and stability (Reetz *et al.*, 2006a).

As for rational design, one of the prerequisites for focused mutagenesis is to be able to identify the regions to mutate. This requires some knowledge of the structural and functional properties of the target enzyme. Information can be gathered using the same techniques mentioned above. Once positions are chosen, the most common approach employed to introduce mutations is based on the principle of whole plasmid amplification with mutagenic primers (Papworth *et al.*, 1996; Zheng *et al.*, 2004). The target gene is cloned into a plasmid which is subsequently amplified with primers that contain the desired modifications (Neylon, 2004).

Mutagenic primers are designed to anneal to the target regions of the gene despite these modifications. During amplification, the modifications are introduced in the resulting vectors, which are then used for enzyme expression. The same technique can be used to introduce all (saturation mutagenesis), or part of, the amino acid pool.

Table 2.1: DNA degeneracy table

Identification Letter	Mixture of nucleobases
T	T
C	C
A	A
G	G
Y	C T
R	A G
S	G C
W	A T
K	T G
M	A C
B	C G T
D	A G T
H	A C T
V	A C G
N	A C G T

Table 2.2: Examples of codon degeneracy

Codon degeneracy	Number of codons generated	Encoded amino acids
NNK	32	Arg, Asn, Asp, Cys, Gln, Glu, Gly, His, Ile, Leu, Lys, Met, Phe, Ser, Trp, Tyr, Val, Pro, Thr, Ala
NDT	12	Arg, Asn, Asp, Cys, Gly, His, Ile, Leu, Phe, Ser, Tyr, Val

To achieve variation, mutagenic primers contain a mixture of nucleobases at the target positions. This is referred to as degeneracy, and is conventionally indicated by letters (Table 2.1). For example, N degeneracy refers to the mixture of all nucleotides (A, T, C, G), and K degeneracy refers to the mixtures of only T and G. By combining degeneracy at three consecutive positions, it is possible to generate mixtures of different codons which encode different amino acids. For example, the NNK degeneracy, a commonly used codon degeneracy (Reetz, 2011; Reetz *et al.*, 2008), consists of a mixture of A, T, C and G, at the first two

positions (N), and of T and G, at the third position (K) (Table 2.1). When used in a mutagenic primer, NNK will give a combination of 32 codons which will encode all 20 amino acids (Table 2.2). Another codon degeneracy is NDT (Reetz, 2011; Reetz *et al.*, 2008). NDT gives a combination of only 12 codons and encodes 12 amino acids (Reetz, 2011; Reetz *et al.*, 2008) (Table 2.2). NNK can be employed to provide a full coverage of the amino acid pool, whereas NDT can be used where less variation is desired. Other levels of variations can be obtained by using different degeneracies. Depending on the level of variation and on the number of positions mutated simultaneously, the number of combinations of encoded amino acid changes drastically. The higher the degeneracy and the number of positions, the higher will be the number of combinations and, consequently, the bigger the screening effort.

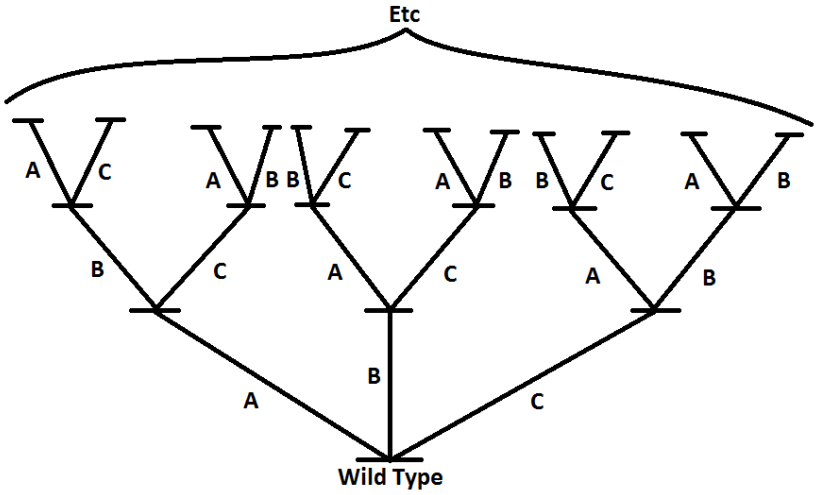


Figure 2.20: Principles of Iterative Saturation Mutagenesis
Schematic illustration involving three randomization sites A, B, C.

Iterative Saturation Mutagenesis (ISM) (Reetz and Carballeira, 2007) provides a successful approach to minimise the screening effort. Several residues are selected but, rather than introducing variation at each position simultaneously, the residues are first split into smaller groups (2-3 amino acids) and saturation, or partial, mutagenesis is then performed at each site. Groups are formed depending on the position of the residues on the enzyme. Residues

in close proximity are usually grouped together, because it is more likely to stimulate synergistic and cooperative effects (Reetz *et al.*, 2010a). The variants generated from mutagenesis at each site are screened, and those that have shown an enhancement for the target property are used as a template for further rounds of mutagenesis at the other sites, aiming to detect further improvements. The process can continue iteratively until the desired grade of enhancement is reached (Figure 2.20).

The power of ISM is the capacity to reduce screening effort because it narrows down the number of residues mutated in each round of mutagenesis. This does not compromise the likelihood to evolve enhanced enzymes because, even if only in a smaller and defined region of protein sequence space, the evolutionary pressure is still applied (Reetz and Sanchis, 2008). Moreover, in each subsequent cycle, the enzyme is allowed to "respond" to the changes at other sites (Horsman *et al.*, 2003; Morley and Kazlauskas, 2005). The benefit of this approach is confirmed by the several examples where ISM was more successful than directed evolution (Bartsch *et al.*, 2008; Bougioukou *et al.*, 2009; Prasad *et al.*, 2011; Reetz *et al.*, 2005; Reetz *et al.*, 2006a; Reetz *et al.*, 2006b; Reetz *et al.*, 2010b; Reetz *et al.*, 2006c; Zheng and Reetz, 2010).

2.2.2.4 Choosing the best approach

Choosing between directed evolution, rational design and semi-rational design depends on several parameters. First of all, the development of an appropriate screening or selection assay, and the evaluation of its throughput capacity, is a bottle neck (Reymond, 2005). If a high throughput assay that detects the decarboxylation of itaconic, mesaconic, citraconic and citramalic acid, or the formation of MAA, can be developed, a very large number of variants could potentially be generated by directed evolution, or by focused mutagenesis with a high level of variations, and then screened or selected. If the throughput level is limited, more

rational approaches must be employed, and ISM has demonstrated to be a powerful tool (Reetz *et al.*, 2008). Nevertheless, since the ability to perform rational or semi-rational design depends on the possibility to identify target residues, ISM will only be possible if enough information on the selected decarboxylases is available. Finally, it is also important to examine cost and time efforts, because mutagenesis procedures and high throughput screenings can be expensive and time-consuming. The final decision can be different, and must be driven by a balanced analysis of advantages and disadvantages, possibilities and impossibilities, on a case-by-case basis.

3 Aims and objectives

The aim of the study was to identify and characterise suitable decarboxylases that could catalyse the formation of MAA from itaconic, mesaconic, citraconic or citramalic acid. The objectives were:

1. Identify candidate enzymes that could potentially lead to the desired conversion.

   Although a few enzymes are characterised by high specificity towards particular substrates, many present instead substrate promiscuity (Humble and Berglund, 2011) that can be used to catalyse alternative reactions or to form unprecedented products (Bornscheuer and Kazlauskas, 2004). Therefore, the objective was to seek for candidate enzymes with potential to promiscuously catalyse the formation of MAA from the substrates of interest. This was done within the large class of decarboxylases (EC 4.1.1.), since these are the enzymes that catalyse similar reactions in nature. The objective was to focus on those decarboxylases that presented similarity between the natural substrates and the substrates of interest, that provided the mechanistic requirements for the target reactions to occur, and that have shown, where possible, some level of substrate promiscuity.

2. Obtain the genes encoding the candidate enzymes.

   The objective was to obtain the genes by amplification from the genomic DNA of original microorganisms with Polymerase Chain Reaction (PCR) or by gene synthesis.

3. Express the selected enzymes to produce soluble, stable and active catalysts for testing.

   The objective was to obtain a high amount of enzyme in the minimum time possible. The system used included cloning

genes into pET-20b(+) vector and subsequently transforming E. coli BL21(DE3)pLys. This system (Appendix 9.1, Page 196) is a cheap and well-established solution which allows, in principle, to overexpress the desired enzyme in a few hours, to an amount of more than 50% of the total cell protein (Novagen, 2003).

4. Test enzymes for activity on itaconic, mesaconic, citraconic and citramalic acid to evaluate their ability to catalyse the target reaction.

   The objective was to test the enzymes avoiding laborious and time-consuming purification procedures. Therefore, tests were done in crude extracts and the reaction conditions employed were based on literature. In order to be sure that the enzymes were expressed in active form, they were always tested on their natural substrates before being tested on the substrates of interest.

5. Further characterise some of the enzymes

   The purpose of this objective was to gain more insights into the potential use of the candidate enzymes for the bioproduction of MAA and to develop future strategies for enzyme selection.

6. Perform protein engineering to improve/generate activity

   Whether the candidate enzymes were active or not, the next objective was to perform protein engineering to improve catalytic efficiency or to generate new variants with the desired activity towards the substrates of interest. In this regard, since there is no know general screening method for decarboxylases, the first task was to develop a robust and reliable assay to enable screening for improved enzyme variants. The next task was to generate mutagenic libraries and to screen for MAA formation.

4 Materials and methods

4.1 Materials

4.1.1 Chemicals

The following chemicals were purchased from Sigma-Aldrich: yeast extract, peptone, glucose, agar, carbenicillin, chloramphenicol, ethanol (EtOH), isopropanol, isobutanol, glycerol, imidazole, isopropyl β-D-1-thiogalactopyranoside (IPTG), β-mercaptoethanol, acrylamide:N,N'-methylenebisacrylamide (37.5:1), sodium dodecyl sulphate (SDS), tetramethylethylenediamine (TEMED), dimethyl sulfoxide (DMSO), ethylenediaminetetracetic acid (EDTA), Tween 80, Triton-X-100, Tris, bromocresol green, cresol red, phenol:chloroform:isoamyl alcohol, magnesium sulphate ($MgSO_4$), 99.9% deuterium oxide (D_2O), HEPES, ammonium acetate, sodium acetate, hydrochloric acid (HCl), sodium chloride (NaCl), magnesium chloride ($MgCl_2$), potassium chloride (KCl), zinc chloride ($ZnCl_2$), potassium permanganate ($KMnO_4$), ammonium molybdate ($(NH_4)_6Mo_7O_{24}$), sodium hydroxide (NaOH), sulphuric acid (H_2SO_4), L-ascorbic acid, citric acid, bromide-bromate solution, sodium carbonate ($NaHCO_3$), soda-lime, *trans*-cinnamic, sorbic acid, itaconic acid, citraconic acid, mesaconic acid, *cis*-aconitic acid, fumaric acid, methacrylic acid (MAA), 3-hydroxyanthranilic acid (3-HAA), 3-hydroxyisovaleric acid (3-HIV), (*R*)-citramalic acid, (*S*)-citramalic acid, isoprenol, (*R*,*S*)-mevalonate-5-diphosphate (MVAPP), (*R*,*S*)-mevalonate-5-phosphate (MVA-5P), (*R*)-mevalonate (MVA), ATP, NADH and phoshoenolpyruvate (PEP).

3-hydroxybutyrate (3-HB) and 3-hydroxypropionate (3-HP) were purchased from Manchester Organics and TCI Chemicals, respectively. Ethidium bromide ($EtBr_2$), isobutene and agarose were purchased from Fisher. Ethene and propene were kindly donated from the School of Chemistry. Purity of all chemicals was between 92% and >99%.

4.1.2 Reagents

Tris/glycine/SDS Premixed Electrophoresis Buffer, EZBlue Stain solution, TAE buffer and Laemmli Sample Buffer were purchased from Bio-Rad. DNA Loading Dye, Gene Ruler™ 1 Kb Plus DNA Ladder and Spectra™ Multicolour Broad Range Protein ladder were purchased from Fermentas. EB buffer and TE buffer were purchased from Qiagen. BugBuster Reagent™ and nucleotides were purchased from Merck-Millipore. SOC medium and cOmplete EDTA-Free Protease Inhibitor were purchased from Novagen and Roche, respectively.

4.1.3 Enzymes

RNase A, lysozyme from chicken egg white and pyruvate kinase/lactic dehydrogenase (PK/LDH) solution were purchased from Sigma-Aldrich. KOD DNA polymerase and rLysozyme™ Solution were purchased from Novagen. Restriction enzymes (FastDigest®) were purchased from Thermo Scientific. T4 DNA Ligase and Benzonase® nuclease were purchased from New England Biolabs and Merck-Millipore, respectively.

4.1.4 Kits

QIAQuick PCR purification Kit, QIAQuick Gel Extraction Kit, QIAPrep MiniPrep and MidiPrep kit were purchased from Qiagen. TransformAid™ Bacterial Transformation Kit was purchased from Fermentas. JumpStart™ RedTaq® DNA Polymerase kit was purchased from Sigma-Aldrich. QuickChange Lightening Site-Directed Mutagenesis Kit was purchased from Agilent and DC™ Protein Assay Kit from Bio-Rad.

4.1.5 Plasmids

pET-20b(+) was purchased from Novagen. pET-20b(+)-NT, a modified version of pET-20b(+) used to allow target proteins to be

expressed with N-terminal His-tag, was constructed in the laboratory. The N-terminal His-tagging sequence was excised from pET-28a(+) using *NdeI* and *XbaI* restriction sites and inserted in pET-20b(+) linearized with the same restriction enzymes. pBMH vector was obtained from Biomatik as a carrier of the synthesised genes.

4.1.6 Strains

Aspergillus terreus, A. niger N402 (Bos *et al.*, 1988) and *Saccharomyces cerevisiae* W303 (Ralser *et al.*, 2012) were kindly donated by Professor David Archer (School of Biology - University of Nottingham). *Escherichia coli* JM107 was purchased from Fermentas-Thermo Scientific (UK) and competent *E. coli* BL21(DE3)pLysS from Merck Millipore (Germany). *E. coli* XL-Gold Ultracompetent cells were obtained as part of the QuickChange Lightning Site-Directed Mutagenesis Kit purchased from Agilent Technologies (US)

4.2 Methods

4.2.1 *In silico* analysis

4.2.1.1 Search for candidate enzymes

Databases such as CAS Scifinder (www.cas.org) and NCBI Pubmed (www.ncbi.nlm.nih.gov) were used to obtain articles, patents, and other scientific publications. Protein sequences, substrate range information and other significant functional data, were collected with the help of four main on-line enzyme and genome databases: the Kyoto Encyclopaedia of Genes and Genomes (www.genome.jp/kegg), the Uniprot database (www.uniprot.org), the Comprehensive Enzyme Information System - BRENDA (www.brenda-enzymes.org) and the RCSB Protein Data Bank (www.rcsb.org).

4.2.1.2 Multiple sequence alignments

Multiple sequence alignments (MSA) of MVD and homologues were generated using MUSCLE (Edgar, 2004) and rendered with ESPripT (Gouet *et al.*, 2003). The organisms from which the amino acid sequences were used and the respective GenBank (Benson *et al.*, 2013) sequence accession numbers were as follows: *Staphylococcus epidermidis* (AAO03959); *Saccharomyces cerevisiae* (AAC49252); *Legionella pneumophila (*AAU28109); *Trypanosoma brucei* (EAN78728); *Candida albicans* (AAF19399); *Homo sapiens* (AAC50440); *Mus musculus* (CAC35731); *Rattus norvegicus* (AAB00192); *Arabidopsis thaliana* (AEE79204); *Ginkgo biloba* (AAV32433); *Staphylococcus aureus* (BAB56753); *Streptococcus pyogenes* (AAK33797); *Streptococcus pneumonia* (CCP34068); *Picrophilus torridus* (AAT43941), *Thermoplasma volcanium* (BAB59465); *Thermoplasma acidophilum* (CAC12426); *Ferroplasma acidarmanus* (AGO61795);

4.2.1.3 Search for homologues and phylogenetic analysis

Sequences were searched with BLASTp (Altschul *et al.*, 1997) using the *P. torridus* MVD amino acid sequence or the newly discovered *T. acidophilum* Ta0762 sequence (GenBank accession numbers AAT43941 and CAC11895, respectively) within non-redundant sequences database (November 2014). For the analysis of *P. torridus* MVD, the cut off was set at a E value of 0.1 and the 47 most significant hits were aligned with MUSCLE (Edgar, 2004). For *T. acidophilum* Ta0762, the cut off was set at 40% sequence identity. Phylogenetic reconstruction for the selected homologues of *P. torridus* MVD was performed with MEGA6 (Tamura *et al.*, 2013) by Maximum Likelihood Test.

4.2.1.4 Structural analysis

For *S. cerevisiae* MVD, structural analysis was done using its available crystal structure (PDBID: 1FI4) (Bonanno *et al.*, 2001). For *P. torridus* MVD and *A. niger* OhbA1, homology models were obtained by sending the amino acid sequences to the I-TASSER

portal (http://zhanglab.ccmb.med.umich.edu/I-TASSER/), an on-line service for protein structure and function prediction. Upon receiving the sequence, i-TASSER identifies templates from the PDB database (www.rcsb.org) and builds a structural model of the target protein by iterative fragment assembly simulation (Roy *et al.*, 2010; Zhang, 2008). To analyse the binding of (*S*)-citramalic acid in *S. cerevisiae* MVD and the binding of *trans*-cinnamic acid in *A. niger* OhbA1, a structure file of the two substrates geometrically optimised were first generated using the WebMO platform (www.webmo.net) and AutoDock Tools (Trott and Olson, 2010), and the ligands were then docked into the respective enzyme using AutoDock Vina, a package based on an iterated local search global algorithm (Trott and Olson, 2010). Water molecules were removed prior to the analysis and rigid docking was performed with standard parameters. To analyse the binding of (*S*)-citramalic acid in *P. torridus* MVD, the homology model of the enzyme was superimposed to *S. epidermidis* MVD co-crystalized with MVAPP (PDBID: 4DU7) (Barta *et al.*, 2012) using the PyMOL Molecular Graphics System. (*S*)-citramalic acid was built on the scaffold of MVAPP using the Build tools of PyMOL. In all cases, PyMOL was also used for visual evaluation of structure and binding.

4.2.2 Growth of strains

4.2.2.1 General handling and preparation of media

A. terreus, *A. niger* N402 and *S. cerevisiae* W303 were grown on YPD, composed of yeast extract (10 g/L), peptone (20 g/L) and glucose (20 g/L). *E. coli* JM107 and *E. coli* XL-Gold were grown on Luria Bertani (LB) medium, composed of peptone (10 g/L), yeast extract (5 g/L) and NaCl (10 g/L). *E. coli* BL21(DE3)pLysS was grown on LB with addition of glucose (10 g/L) and chloramphenicol (34 mg/L). When *E. coli* strains were engineered with pET-20b(+) or pET-20b(+)-NT, carbenicillin (50 mg/L) was also added to the media.

Media were dissolved in distilled water (dH_2O) and then autoclaved. For the preparation of agar plates, agar (15 g/L) was also added before autoclaving. Agar media were then cooled to ~50°C and poured manually in Petri dishes. Carbenicillin, chloramphenicol and glucose were added separately as required from filter sterilised 50 g/L stock solution prepared in dH_2O, 34 g/L stock solution prepared in 100% EtOH, and 200 g/L stock solution prepared in dH_2O, respectively.

For long-term storage of all organisms, fresh liquid medium (5 mL) was inoculated and incubated (30°C for *Aspergillus sp.* and *S. cerevisiae*, 37°C for *E. coli*) until the exponential growth phase was reached. Samples (900 µL) were transferred into cryo-vials together with sterile 80% glycerol (100 µL) and stored at -80°C. For experiments inocula, strains were grown freshly overnight on agar and single colonies were transferred to fresh liquid medium as required using a sterile plastic loop. For *Aspergillus sp.*, spores were collected by adding 0.01% Tween 80 (2 mL) to the plates and then transferred into fresh liquid medium.

4.2.2.2 Expression of enzymes

For optimisation of enzyme expression and for the preparation of cultures for enzymatic assays, *E. coli* BL21(DE3)pLysS transformed with pET-20b(+) or pET-20b(+)-NT containing the genes of interest (20 mL) were incubated at 37°C with 200 rpm shaking. Each time, a control culture of *E. coli* transformed with empty plasmid was also prepared. When OD_{600} reached approximately 1, the cultures were collected by centrifugation for 20 min at 5000 X *g*, resuspended in fresh medium (5 mL), and used as the inoculum for a bigger culture (100 mL when optimizing enzyme expression and 400 mL when preparing cultures for enzymatic assays). When OD_{600} reached approximately 0.6, expression of the enzymes was induced by addition of filter sterilized IPTG to a final concentration of 0.4 mM. Temperatures and incubation times are described in the text

For expression of the enzymes in 96-well plates, single colonies of *E. coli* BL21(DE3)pLysS transformed with pET-20b(+) or pET-20b(+)-NT containing the genes of interest, or obtained after a fresh transformation with mutagenic libraries, were picked using a sterilised cocktail stick and individually inoculated into 96-deep well plates containing fresh medium (0.4 mL each well). Single colonies of *E. coli* transformed with empty plasmid were also prepared and inoculated as control. Plates were incubated overnight at 37°C with 250 rpm shaking. The cultures where diluted by discarding (wild type enzymes) or by transferring (enzyme variants) 0.3 mL to a 96-well cryo-plate containing 80% glycerol (75 uL) for storage at -80°C, and adding fresh medium (0.8 mL) to the remaining cultures (0.1 mL).The plates were incubated at 37°C. After 6-8 h, IPTG (90 µL) was added to each well at a final concentration of 0.4 mM and plates were incubated overnight at 37°C (14-16 h).

4.2.3 Molecular biology

4.2.3.1 Oligonucleotides

The oligonucleotides used for gene amplification (Table 4.1), to perform mutagenesis (Table 4.2) or for gene sequencing (Table 4.3), were synthesised by Integrated DNA Technology (US).

Table 4.1: Synthetic oligonucleotides for gene amplification[a]

Name	Sequence
PR1	5'-AAACATATGCTCCTATTTCCAAGAAGAACTAATATAGCCTTT- 3'
PR2	5'-AAAGCGGCCGCTTACTTGCTTTTTATTCCTT- 3'
PR3	5'-AAACATATGAGGAAGCTAAATCCAGCTTTAGAATTTAGAGACTTTATC- 3'
PR4	5'-AAAGCGGCCGCTTATTTATATCCGTACCTT- 3'
PR5	5'-AAACATATGTTCAACTCACTTCTATCCGGTACTACTACACCAAA- 3'
PR6	5'-AAAGCGGCCGCTTATTTTTCCCATCCATTC- 3'
PR7	5'-AAACATATGTCTGCGCAACCTGCTCACCTG- 3'
PR8	5'-AAAGCGGCCGCTTAGTTGCTGAAACCCATC- 3'
PR9	5'-GCCGCTTGACGGATGAAACGCACACCATGATCGGCTCTC- 3'
PR10	5'-GCACACCATGATCGGCTCTCTGGCT- 3'
PR11	5'-AAACATATGACCAAGCAATCTGCGGACAGCAAC- 3'
PR12	5'-AAAGCGGCCGCTTATACCAGTGGCGATTTC- 3'
PR13	5'-TGCTGCAACAGGCCCCAGTTTCTGTCCATATCCAATCACCCTGCAGGCC- 3'
PR14	5'-AAACTGGGGCCTGTTGCAGCAGCCATG- 3'

[a] PR1-PR2 are for amplification of the gene encoding *S. cerevisiae* PAD1, PR3-PR4 for the gene encoding *S. cerevisiae* FDC1, PR5-PR6 for the gene encoding of *A. niger* PadA1, PR7-PR8-PR9-PR10 for the gene encoding *A. niger* OhbA1, PR11-PR12-PR13-PR14 for the gene encoding *A. terreus* CadA1.

4.2.3.2 Extraction of genomic DNA

Protocols were kindly advised by Lee Sunburne (School of Life Sciences, University of Nottingham). *S. cerevisiae* W303 was grown overnight in fresh medium (100 mL) at 30°C with shaking at 200 rpm. A sample (2 mL) was centrifuged at 12000 X *g* for 5 min, the pellet was washed in dH_2O (1 mL), collected again by centrifugation and resuspended in 10 mM Tris (pH 8), 100 mM NaCl, 1 mM EDTA, Triton-X-100 (2% v/v), SDS (2% w/v) (200 µL). Phenol:chloroform:isoamyl alcohol (25:24:1 v/v, 200 µL) was added and cells were lysed by four agitation cycles in the presence of glass beads (30 s agitation, 30 s on ice). The resulting liquid was transferred into a heavy phase lock tube (5prime). TE buffer (200 µL) was added and the solution vortexed briefly and centrifuged at 12000 X *g*. The aqueous layer was transferred into a clean tube. DNA was precipitated by adding ice-cold EtOH (1 mL) and by incubating at -20°C for 30 min. The precipitate was collected by centrifuging at 12000 X *g* and then resuspended in TE buffer (400 µL). Ammonium acetate buffer (4 M, pH 4.5, 10 µL) and EtOH (1mL) were then added and the liquid transferred into a clean tube. After 1 h at -20°C and 20 min centrifugation at 12000 X *g*, the pellet was collected and resuspended in TE buffer (100 µL).

A. terreus and *A. niger* N402 were grown at 30°C in fresh medium (100 mL) until the mycelium developed (1-2 d). The mycelium was harvested through Miracloth (EMD Millipore), washed three times with dH_2O and dried with paper towels. After overnight freeze-drying, the mycelium was ground with mortar and pestle. At this stage liquid nitrogen was added continuously to keep the mycelium frozen while grinding it. Ground mycelium (200 mg) was resuspended in 50 mM Tris-HCl (pH 7.5), 10 mM EDTA, 50 mM NaCl, SDS (1% w/v) (800 µL) and placed in a heavy phase lock tube (5prime). The mixture was incubated at 65°C for 30 min and mixed frequently. After cooling down the sample on ice for 5 min, phenol:chloroform:isoamyl alcohol (25:24:1 v/v, 800 µL) was added, the tube was shaken vigorously for 30 s and subsequently

centrifuged at 12000 X *g* for 10 min. The aqueous phase was transferred into a clean tube and the DNA extracted as described above. In both cases, treatment with RNase A (1 µL) was performed at 37°C for 5 min.

4.2.3.3 Polymerase chain reactions

Polymerase chain reaction (PCR) was performed to amplify the genes of interest. The reactions (50 µL) contained genomic DNA at different concentrations, primers (0.4 mM), nucleotides (0.3 mM), Novagen Buffer for KOD polymerase and KOD DNA Polymerase (2.5 U). When required, DMSO and $MgCl_2$ were added at different concentrations. Forward and reverse primers were designed with *NdeI* and *NotI* restriction sites, respectively, to allow subsequent gene cloning procedures (see section 4.2.3.5).

After transformation of *E. coli* cells with pET-20b(+) and pET-20b(+)-NT plasmids, colony PCR was performed to confirm the presence of the genes. The reactions (50 µL) contained JumpStart REDTaq® PCR Buffer, nucleotides (0.2 mM), T7 promoter and T7 terminator primers (0.1 mM), JumpStart™ RedTaq® DNA Polymerase (2.5 U). Several colonies were randomly transferred into the reaction mixtures using a sterile 100 µL pipette tip. The following cycling conditions were used: 94°C for 3 min, 35 x (94°C for 30 s; 50°C for 30 s; 72°C x 1 min for every kb of expected product) and 72°C for 5 min.

Table 4.2: Synthetic oligonucleotides for mutagenic amplification[a]

Name	Sequence
PR15	5'-CGACCATTGGTATTATTCTGCTGNDKGGCNDTAGCGACAAAAAAAACCGTATTCC- 3'
PR16	5'-GGAATACGGTTTTTTTTGTCGCTAHNGCCMHNCAGCAGAATAATACCAATGGTCG- 3'
PR17	5'-CGCAAACCGAGCAACGAANDTCACGAANDKATCATCAAACATGAGAACTATCG- 3'
PR18	5'-CGATAGTTCTCATGTTTGATGATMHNTTCGTGAHNTTCGTTGCTCGGTTTGCG- 3'
PR19	5'-GATTTCTCGTATCGCGCGCAAGGGCTCTNDKNDTGCGTGCCGTTCTCTGTTTG- 3'
PR20	5'-CAAACAGAGAACGGCACGCAHNMHNAGAGCCCTTGCGCGCGATACGAGAAATC- 3'
PR21	5'-GACATCAAGAAGGACGTTTCTNDTACCCAGGGTNDKCAGCTGACCGTGGC- 3'
PR22	5'-GCCACGGTCAGCTGMHNACCCTGGGTAHNAGAAACGTCCTTCTTGATGTC- 3'
PR23	5'-ATCATGCCCATGTCTTCTGCCGGCCGCTTGACGGAT- 3'
PR24	5'-ATCCGTCAAGCGGCCGGCAGAAGACATGGGCATGAT- 3'

[a] PR15-PR16 and PR17-PR18 were used for generation of *P. torridus* MVD library A and B, respectively. PR19-PR20 and PR21-PR22 P were used for generation of *S. cerevisiae* MVD library A and B, respectively. PR23-PR24 were used for generation of the *A. niger* OhbA1 Cys316Ala mutant

For mutagenesis, PCR amplification of plasmids was performed according to the QuickChange Lightening Site-Directed Mutagenesis Kit. Reaction mixtures (50 µL) contained QuickChange Lightning reaction buffer, template plasmid (0.5 – 2 ng/µL), primers (Table 4.2) (2.5 ng/µL), nucleotides mixture (1 µL), QuickSolution reagent (1.5 µL) and QuickChange Lightning enzyme (2.5 U/µL). The following cycling conditions were used: 95 °C for 2 min, 18 x (95 °C for 20 s, 60 °C for 10 s and 68 °C for 1 min) and 68 °C for 5 min. After cooling down the reaction on ice for 2 minutes, *DpnI* (10 U/µL) was added and the reactions incubated at 37 °C for 15 min. *E. coli* XL-Gold was then transformed with the resulting mixtures (2 µL). For the genes encoding *P. torridus* and *S. cerevisiae* MVD, saturation mutagenesis was performed. For the gene encoding *A. niger* OhbA1, alanine-mutagenesis was performed. In both cases, pET-20b(+) or pET-20b(+)-NT containing the target genes were first extracted and purified from *E. coli* JM107, and then used as a template.

4.2.3.4 Gene synthesis

Some of the selected genes of interest were synthesised by BioMatik, optimised for overexpression in *E. coli* and with the addition of *NdeI* and *NotI* restriction sites. The genes were delivered in lyophilised form as inserts in the pBMH plasmid. Plasmids were reconstituted in EB buffer and used for further cloning procedures.

4.2.3.5 Preparation of plasmids

The genes of interest amplified by PCR were purified using the QIAQuick PCR purification Kit and then digested with NdeI and NotI. The genes of interest synthesised in pBMH vectors were obtained by digesting the plasmids. Digestion reactions (50 µL) contained PCR product or plasmid (30 µL), FastDigest® buffer, restriction enzymes (1 µL) and nuclease free water. Reactions were incubated at 37°C for 5 h. All fragments were run on an agarose gel and purified using the QIAQuick Gel Extraction Kit. The fragments were then ligated in pET-20b(+) or pET-20b(+)-NT pre-linearised with the same

restriction enzymes. Ligation mixtures (20 µL) contained T4 DNA ligase, linearized plasmid (50 ng) and inserts (3:1 molar ratio with plasmids). The reaction was performed at room temperature for 30 min.

4.2.3.6 Cell transformation

Transformation of *E. coli* JM107 was performed using the TransformAid™ Bacterial Transformation Kit. Transformation of *E. coli* BL21(DE3)pLysS and *E. coli* XL-Gold was performed by heat shock. The purchased competent cells were conserved in aliquots (20 µL) at -80°C. On the day of transformation, the required amount of aliquots was thawed on ice for 2-5 min and the plasmids (10 ng), either purified from *E. coli* JM107 or generated with the QuickChange Lightening Site-Directed Mutagenesis Kit, directly added into the tubes. After gentle stirring, the mixtures were incubated on ice for 5 min, heated for exactly 30 s in a 42°C water bath, and placed back on ice for 2 min. Room temperature SOC medium (80 µL) was then added and the cells were incubated at 37°C for 60 min with 250 rpm shaking. The cultures were finally plated on agar.

4.2.3.7 Extraction and purification of plasmids

E. coli was sub-cultured in agar medium and single colonies inoculated individually in liquid medium (5-100 mL). After overnight growth at 37°C, the plasmids were extracted and purified using the QIAPrep MiniPrep or MidiPrep Kit.

4.2.4 Isolation and purification of enzymes

4.2.4.1 Preparation of soluble and insoluble fractions for PAGE analysis

E. coli BL21(DE3) pLysS cells expressing the enzymes of interest were sampled at intervals. Samples (1.5 mL) were centrifuged at 8000 X *g* for 10 min and pellets were resuspended in the 100 mM potassium phosphate buffer (pH 8), BugBuster Reagent™, cOmplete

EDTA-Free Protease Inhibitor (1 tablet every 50 mL) and Benzonase® (2.5 U/mL). The amount of buffer used for each sample was calculated depending on the OD which the original sample was taken at, using the formula: $\text{Solution amount} = (OD/(0.2) \times 45\mu L)) \,/\, 2$. This normalised the total protein concentration among the samples. Resuspended cells were incubated for 15 min at room temperature with gentle shaking and an amount of liquid equal to the lowest quantity of BugBuster Reagent™ used for cell resuspension was transferred into a fresh tube. Tubes were centrifuged at 12000 X *g* for 20 min at 4°C and the supernatant (20 µL) was transferred into another fresh tube. This was used as soluble fraction sample. The remaining pellet was resuspended into 100 mM potassium phosphate buffer (pH 8), using the same buffer volume as the volume of liquid transferred before centrifugation. A sample (20 µL) was then transferred into a fresh centrifuge tube and used as insoluble fraction sample. The fractions were then analysed by Polyacrylamide Gel Electrophoresis (PAGE).

4.2.4.2 Preparation of crude extract for enzyme assays

E. coli BL21(DE3) pLysS cells containing the enzymes of interest were collected by centrifugation. For cells grown in flasks, centrifugation was performed at 12000 X *g* for 20 min and for cells grown in 96-well plates at 3000 X *g* for 30 min. Cells were resuspended in buffer with addition of nuclease and protease inhibitor as described above. For cells grown in flasks, the cell resuspension (50 mL) was then lysed with a Constant Systems cell disruption system, cell debris was removed by centrifugation at 18000 X *g* for 15 min, and crude extract transferred in fresh tubes. For cells grown into 96-well plates, the resuspension solutions (100 µL each well) also contained either lysozyme from chicken egg white (1 mg/mL), rLysozyme™ Solution (7.5 KU/mL) or BugBuster® Reagent, and lysis was performed by incubation for 30 min at 37°C, 25°C and room temperature, respectively. After lysis, buffer (200 µL) was added to each well to dilute to a total protein concentration of 0.5-1 mg/mL, plates were centrifuged again at

3000 X *g* for 15 min at 4°C to remove cell debris, and crude extract transferred into a fresh 96-well plate. In all cases, the buffer used was the same as that used for the enzymatic assays, unless differently indicated.

4.2.4.3 Preparation of purified enzymes

Crude extracts containing his-tagged *P. torridus* MVD, *S. cerevisiae* MVD, *P. fluorescens* ACMSD and *A. niger* OhbA1 were prepared in 20 mM sodium phosphate buffer (pH 7.4) with addition of 30 mM imidazole and 500 mM NaCl. The enzymes were then purified using an AKTA Protein Purification System. Crude extracts (10 mL) were applied to a 1 mL- HisTrap™ FF Crude column (GE Healthcare), proteins were eluted with a linear gradient of 20 mM sodium phosphate buffer (pH 7.4) containing 500 mM imidazole and 500 mM NaCl (1 mL/min) at 4°C, and elution was monitored by UV (280 nm). The fractions (2 mL) containing the enzymes were collected and combined. The buffer was then replaced with the buffer in which the enzyme assays were performed, by doing several concentration/dilution cycles using a Vivaspin Sample Concentrator. Fresh enzymes were used each time.

4.2.5 Enzyme assays

4.2.5.1 Phenylacrylic acid decarboxylases

S. cerevisiae and *A. niger* phenylacrylic acid decarboxylases in crude extract were tested either as a mixture (1:1 ratio) of PAD1 and FDC1, or PadA1 and OhbA1, respectively, or as individual enzymes. The reaction mixtures (1 mL) contained 50 mM sodium phosphate buffer (pH 6.5), crude extracts (1 mg/mL total protein) and substrate (*trans*-cinnamic, sorbic, itaconic, citraconic or mesaconic acid) at the concentration described in the text.

Purified *A. niger* OhbA1 was tested for activity on *trans*-cinnamic acid. The reaction mixtures (1 mL) contained *A. niger* OhbA1 (100 µg), 50 mM sodium phosphate buffer (pH 6.5) and

trans-cinnamic acid (1 mM). The assay was also performed with the addition of crude extract (100 µL) from *E. coli* BL21(DE3) pLysS, and with the addition of filtrate and retained solution (100 µL) of the same crude extract after passing it through a Vivaspin Sample Concentrator with a 2 kDa molecular cut off. $CuSO_4$, $ZnCl_2$, $MgCl_2$, $FeSO_4$ (all at 1 mM), ATP (5 µM), NADH, FAD, FADH (all at 10 µM), dithiothreitol, 2-mercaptoethanol, homocysteine and glutathione (all at 1 mM) were also added as required.

The assays were performed in sealed GC vials (2 mL nominal volume) with polytetrafluoroethylene (PTFE) screw caps and incubated at 30°C for 24 h with gentle shaking. Substrate consumption and product formation were monitored at intervals either by Gas Chromatography-Mass Spectrometry (GC-MS) or High Pressure Liquid Chromatography (HPLC). 1M H_2SO_4 (20 µL) was added to the samples (1 mL) to stop the enzymatic reaction. For the detection of styrene and 1,3-pentadiene, samples (1 mL) were then extracted with diethyl ether and hexane (1:1), respectively. The organic phase was dried with $MgSO_4$ and transferred into GC vials which had PTFE screw caps before being analysed. Standards treated the same way did not decompose.

4.2.5.2 *cis*-Aconitate decarboxylase

A. terreus cis-aconitate decarboxylase was tested in crude extract similarly to previously described (Bentley and Thiessen, 1957; Dwiarti *et al.*, 2002). The reaction mixtures (1 mL) contained 0.2 M sodium phosphate buffer (pH 6.2), crude extract (1 mg/mL total protein) and substrate (*cis*-aconitic acid, mesaconic acid or citraconic acid) at the concentration described in the text. The assays were performed in microcentrifuge tubes and incubated at 45°C for 24 h with gentle shaking. Substrate consumption and product formation were monitored at intervals by HPLC. Enzymatic reactions were stopped as described above.

4.2.5.3 *α*-Amino-*β*-carboxymuconate-*ε*-semialdehyde decarboxylase

P. fluorescens *α*-amino-*β*-carboxymuconate-*ε*-semialdehyde decarboxylase (ACMSD) was tested in crude extract and purified form. *P. fluorescens* HAO crude extract was used to produce ACMS from 3-hydroxyanthranilic acid (3-HAA), as described previously (Huo *et al.*, 2013; Li *et al.*, 2006; Muraki *et al.*, 2003). 3-HAA (0.4 M) was freshly prepared in DMSO and then diluted 100 fold in air-saturated 25 mM HEPES buffer (pH 7) with addition of 5% glycerol. Reaction mixtures (1 mL) were assembled in 1 mL-PMMA cuvettes (BRAND) and contained 25 mM HEPES buffer (pH 7), 5% glycerol, HAO crude extract (1 mg/mL total protein) and 3-HAA at different concentrations. The reaction was performed at 25°C. ACMS production was monitored by measuring the increase in absorbance at 360 nm every 20 s using an Agilent 8453 UV-visible spectrophotometer. Product concentration was calculated using a molar extinction coefficient of 47500 $M^{-1}\cdot cm^{-1}$ (Koontz and Shiman, 1976) when the expected concentration was equal to or lower than 20 µM, and at 320 nm using a molar extinction coefficient of 9600 $M^{-1}\cdot cm^{-1}$ when the expected concentration was higher than 20 µM (Huo *et al.*, 2013).

When ACMS formation had ended (when a plateau was reached in absorbance), ACMSD crude extract (1 mg/mL total protein), or purified ACMSD (5-100 µg/mL), was added to the reaction and the decrease in absorbance at 320 or 360 nm, related to the consumption of ACMS, was monitored. Rates were calculated and subtracted from those of control reactions containing ACMSD-free crude extract. For kinetic characterisation, ACMS was used at a concentration range of 0-80 µM. When testing inhibition, mesaconic (0-4 mM), itaconic, citraconic or fumaric acid (all at 4 mM) were also added to the reaction. Maximum velocities (V_{max}) and Michaelis constants (K_m) were estimated with the equation $v_0=V_{max}[S]/(K_m+[S])$. Inhibition constants (K_i) were estimated with the equation $v_0=V_{max}[S]/K_m(1+[I]/K_i)+[S]$. In all cases, the

non-linear regression models of GraphPad Prism version 6.04 were used.

P. fluorescens ACMSD was also tested for activity on mesaconic, citraconic acid, itaconic and fumaric acid. The reaction mixtures (1 mL) contained 25 mM HEPES buffer (pH 7), 5% glycerol, ACMSD crude extracts (1 mg/mL total protein) or purified ACMSD (100 µg/mL) and substrate at different concentrations. The assays were performed in microcentrifuge tubes and incubated for 24 h at 25°C with gentle shaking. When required, the crude extract or the purified enzyme was incubated with $ZnCl_2$ for 5 h at 4°C prior to start the assay. Substrate consumption and product formation were monitored at intervals by HPLC. Enzymatic reactions were stopped as described above.

4.2.5.4 Mevalonate diphosphate decarboxylases

P. torridus and *S. cerevisiae* mevalonate diphosphate decarboxylases were tested in crude extracts and purified form. Assays were performed at 30-50°C with gentle shaking. Reaction mixtures contained 50 mM Tris-HCl buffer (pH 7-7.5) or 50 mM acetate buffer (pH 5.5), 10 mM $MgCl_2$, 20 mM KCl, substrate at different concentrations and either crude extract (0.5-1 mg/mL total protein) or purified enzyme at the concentration described in the text.

In some experiments, activity was quantified by coupling the release of ADP with NADH oxidation by pyruvate kinase (PK)/lactate dehydrogenase (LDH) similarly to previously reported (Alvear *et al.*, 1982; Barta *et al.*, 2012; Barta *et al.*, 2011; Krepkiy and Miziorko, 2005; Krepkiy and Miziorko, 2004). 0.2 mM NADH, 0.5 mM phosphoenolpyruvate (PEP), 5 mM ATP, PK (16 U/mL), LDH (23 U/mL) were added, and reactions were assembled in PMMA cuvettes (BRAND) (1 mL) or 96-well plates (200 µL). NADH oxidation was measured at 340 nm at 10 s intervals using an Agilent 8453 UV-visible spectrophotometer or a FLUOstar Optima plate reader (BMG Labtech), respectively. NADH concentrations were

calculated using an extinction coefficient of 6220 $M^{-1} \cdot cm^{-1}$ and rates defined as the production of 1 µmol NADH per 1 min. (*R*,*S*)-mevalonate-5-diphosphate (MVAPP), (*R*,*S*)-mevalonate-5-phosphate (MVA-5P), (*R*)-mevalonate (MVA), 3-hydroxyisovalerate (3-HIV), 3-hydroxyisobutyrate (3-HB) and 3-hydroxypropionate (3-HP) were tested as substrate at 500 µM. Kinetic characterisation was performed for MVA, MVAPP, 3-HIV and 3-HB and the concentration range was at 75-900 µM, 50-300 µM and 1-25 mM, respectively. Kinetic analysis was also performed for the co-substrate ATP. MVA and MVAPP were kept in excess (2 mM) and ATP in a range between 18.75 and 300 µM. Maximum velocities (V_{max}) and Michaelis constants (K_m) were estimated as described above. The turnover number (K_{cat}) was calculated dividing V_{max} by the total concentration of enzyme.

Activity was also quantified by monitoring of substrate/product consumption/formation. Reactions were performed in glass vials sealed with PTFE screw caps (1-10 mL) or in 96-well plates (250 µL), and substrates (ATP, 3-HIV, 3-HB, 3-HP, (*R*)-citramalic acid or (*S*)-citramalic acid) were added at different concentrations. Substrate consumption and product formation were monitored at intervals either by HPLC or GC-MS. For HPLC analysis, enzymatic reactions were stopped as described above. For GC-MS analysis of isobutene, ethene and propene, samples were taken in the headspace without stopping the enzymatic reactions. For the detection of isoprenol, samples (250 µL) were extracted with ethyl acetate (750 µL) and the organic phase dried as described above before being analysed by GC-MS. To identify the reaction product of *P. torridus* MVD using MVA as substrate, reaction mixtures were prepared as described previously (Vinokur *et al.*, 2014a) and product analysed by Electrospray Ionization – Mass Spectrometry (ESI-MS) and Nuclear Magnetic Resonance (NMR).

P. torridus and *S. cerevisiae* mevalonate diphosphate decarboxylases were also tested for isobutene formation using whole cells. Cells were grown and induced for 5 h, collected by

centrifugation at 18000 X *g* or 15 min and resuspended in fresh medium containing IPTG. Reaction mixtures (10 mL), containing harvested cells (20 g/L) and 3-HIV, were incubated at 37°C with shaking at 200 rpm in 30 mL-glass vials closed with gas-tight caps with PTFE/Silicone septa. Isobutene formation was monitored by GC-MS in the headspace.

4.2.5.5 Microresp™ assay

Microresp™ was used to detect activity of *A. terreus* CadA1, *P. torridus* MVD, *S. cerevisiae* MVD and MVD variants in crude extract. For *A. terreus* CadA1, crude extract was prepared from cells grown in flasks, and was then transferred into a fresh 96-deep well plate. For wild type MVDs and mutated variants, crude extracts were instead directly prepared from cells grown in 96-deep well plates. Reaction mixtures (0.5 mL) were set up as described above. Substrates (cis-aconitic acid, ATP, 3-HIV and (*S*)-citramalic acid) were added at concentrations described in the text.

The 96-deep well plate containing the reaction mixtures was covered with the seal provided by the manufacturer and closed with a pre-prepared detection plate placed upside down on top of the seal. This was done immediately after addition of substrates to minimize CO_2 loss. Everything was secured with a metal clamp. For preparation of the detection plates, a solution containing 2.5 mM $NaHCO_3$, 150 mM KCl and 12 µg/mL cresol red pH indicator was mixed with 3% agar solution (3:1 ratio) at 60°C. The mixture was then quickly dispensed into appropriate 96-well plates and allowed to solidify. Detection plates were stored in a desiccator containing soda-lime for CO_2 removal and a beaker of water to maintain moisture for 2-3 days before utilization. Plates were incubated with 250 rpm shaking.

Evaluation of CO_2 release was performed using a Decolouring Index (DI). The DI was given by the ratio of the final (A_f) and initial (A_i) absorbance of the detection plates, measured at 570 nm using a FLUOstar Optima plate reader (BMG Labtech).

4.2.6 Analytical methods

4.2.6.1 Estimation of DNA concentration

Plasmid concentration was determined using a NanoDrop ND1000 Spectrophotometer. PCR fragment concentration was determined using agarose gel electrophoresis and by visually comparing bands with standards at known concentrations.

4.2.6.2 Gene sequencing

pET-20b(+) and pET-20b(+)-NT plasmids containing the genes of interest (20 µL), extracted and purified from *E. coli* cells, were diluted to a concentration of 75 ng/µL and shipped to Eurofins MWG Operon (Germany). The T7 promoter and T7 terminator primers (Table 4.3) were used for sequencing.

Table 4.3: Synthetic oligonucleotides used for gene sequencing

T7 promoter	5'-TAATACGACTCACTATAGGG- 3'
T7 terminator	5'-GCTAGTTATTGCTCAGCGG- 3'

4.2.6.3 Agarose gel electrophoresis

Agarose gel electrophoresis was used to purify PCR products and digested DNA fragments, to evaluate results of colony PCRs and to estimate DNA concentration, using a Green Range™ Horizontal Gel Electrophoresis Unit (Scie-Plas). The gel was prepared with TAE buffer (Bio-Rad) containing agarose (1%). After dissolving the agarose by microwaving the solution, ethidium bromide ($EtBr_2$) (1% w/v) was added and the solution poured in the pre-assembled electrophoresis system. The comb was inserted immediately. When solidification was completed, the comb was removed and the DNA samples were loaded in the wells. Before loading, DNA samples were pre-mixed with 6X Loading Dye. Gene Ruler™ 1 Kb Plus DNA Ladder was also added as standard for each analysis. The agarose gel electrophoresis analyses were performed at 50-80 mV for 40-60 min using TAE buffer. Pictures of the gels were obtained an uGenius^3 Gel Documentation System (Syngene).

4.2.6.4 Quantification of total protein in crude enzyme preparations

Protein concentrations were measured using the DC™ Protein Assay Kit. Samples were diluted 10 and 50 times and then transferred (20 µL) into a microcentrifuge tube. Working Reagent A' (100 µL), composed of Reagent S (20 µL) and Reagent A (2 mL), was added and the sample mixed. Reagent B (800 µL) was added into each tube and mixed immediately. After incubation (15 min) at room temperature, absorbance at 750 nm was analysed using an Agilent 8453 UV-Vis spectrophotometer. Protein concentration was calculated based on freshly prepared calibration curves of bovine serum albumin protein standard in a range of 0.2-1.5 mg/mL.

4.2.6.5 Quantification of purified enzymes

The concentration of purified enzymes was calculated from A_{280} measurements with a NanoDrop ND1000 spectrophotometer using extinction coefficients computationally predicted with ProtParam (ExPASy) (Gasteiger *et al.*, 2005). For *P. fluorescens* ACMSD, the coefficient was 50210 $M^{-1}\cdot cm^{-1}$. For *P. torridus* MVD and *S. cerevisiae* MVD, the coefficients were 22015 and 53650 $M^{-1}\cdot cm^{-1}$, respectively. For *A. niger* OhbA1, the coefficient was 69245 $M^{-1}\cdot cm^{-1}$.

4.2.6.6 Polyacrylamide Gel Electrophoresis

SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE) was used to analyse the soluble and insoluble fractions of the enzyme during optimisation of the expression, to analyse fractions during enzyme purification, and to evaluate enzyme expression before enzymatic assays.

Analyses were performed using a Mini-PROTEAN Tetra Electrophoresis System (Bio-Rad). Resolving gel (10 mL) was poured in the pre-assembled electrophoresis system and overlaid with water-saturated isobutanol to avoid air bubble formation. The resolving gel (pH 8.8) contained 0.4 M Tris, 0.13% SDS and acrylamide:N,N'-methylenebisacrylamide (37.5:1) water solution at

different concentrations depending upon the size of the protein to be separated. Acrylamide polymerization was induced by adding TEMED (10 µL every 30 mL of gel) and 10% ammonium persulfate (225 µL every 30 mL of gel). When acrylamide polymerization was completed, water-saturated isobutanol was removed and the resolving gel washed with dH2O. Stacking gel (2 mL) was poured onto the resolving gel and the comb was inserted immediately. The stacking gel (pH 6.8) contained 0.065 M Tris, 0.13% SDS and 4% acrylamide:N,N'-methylenebisacrylamide (37.5:1) water solution. When acrylamide polymerization, induced as described above, was completed, the comb was removed and the wells washed with dH_2O before loading protein samples. Protein samples (20 µL) were mixed with commercial Laemmli Sample Buffer (20 µL) with addition of β-mercaptoethanol (5%). They were then boiled for 5 min at 105 °C before being loaded onto the gel. Spectra™ Multicolour Broad Range was used as the Protein ladder. Runs were performed at 100-300 mV for 1.5-2.5 h using Tris/glycine/SDS Premixed Electrophoresis Buffer (Bio-Rad). "Any kD" and "12%" Mini-PROTEAN TGX Precast Gels were sometimes used instead of preparing gels manually. The gels were subsequently stained overnight with EZBlue Stain (Bio-Rad), washed several times with dH_2O and evaluated by visual observation. Pictures of the gels were obtained an uGenius^3 Gel Documentation System (Syngene).

4.2.6.7 High Pressure Liquid Chromatography

Itaconic, mesaconic, citraconic, citramalic, fumaric, methacrylic and *cis*-aconitic acid were detected and quantified using an Agilent 1200 Series HPLC System. The method was optimised starting from protocols already in use in the lab. The mobile phase was 100% H_2SO_4 (0.005 M, 0.7 mL/min) and components were separated using an Aminex HPX-87H column (Bio-Rad) at 60 °C. Detection was performed at 215 nm using a G1314B VWD (Variable Wavelength Detector) and by Refractive Index using a G1362A Detector at 35 °C. Calibration for each compound analysed was performed using authentic standards at a concentration range between 0.1 and

10 mM (Appendix 9.4, Pages 217-220). Samples were filtered through a 0.2 µm filters and filtrates (20 µL) injected into the apparatus using an autosampler. Data analysis was performed with ChemStation software.

4.2.6.8 Gas Chromatography – Mass Spectrometry

Styrene, 1,3-pentadiene, isoprenol, isobutene, propene and ethene were detected and quantified using an Agilent 7890 A GC-MS System equipped with an Agilent capillary HP5-MS column (0.32 mm x 30 m x 0.5 µm), an Electron Impact ionisation source and a 5975C inert MSD with quadrupole mass analyser. For each compound, detection methods were optimised starting from a standard method already in use in the lab for the detection of similar compounds. Helium was used as carrier gas at 1.2 mL/min. Data analysis was performed using MSD ChemStation software and by comparison with the mass spectra and retention times of authentic standards.

For detection of styrene, 1,3-pentadiene and isoprenol, samples (1 µL) were injected using an autosampler a split ratio of 100:1. For styrene and isoprenol, the inlet temperature was 250°C and, for 1,3-pentadiene, was 150°C. In all cases, the oven temperature was initially kept at 45°C for 5 min but, for styrene and isoprenol, it was increased to 300°C at 20°C/min and, for 1,3-pentadiene, to 200°C at 10°C/min. The final temperature was held for 10 min.

For isobutene, propene and ethene, the headspace (100 µL) of samples was manually injected at an inlet temperature of 150°C and at a split ratio of 100:1. The oven temperature was kept at 40°C for 5 min, increased to 200°C at 20°C/min and then held for 5 min. To improve sensitivity for the detection of 1,3-pentadiene, isobutene, propene and ethene at low concentration, the most characteristic ions were monitored in Selected Ion Monitoring (SIM) mode. For 1,3-pentadiene, ions were 67 and 53 m/z; for isobutene, 39, 41 and 56 m/z; for propene, 39, 41 and 42 m/z; for ethene, 26 and 27 m/z.

A linear calibration of all compounds was obtained using authentic standards in a concentration range of 1 – 7 mM, 0.1 – 10 mM, 0.1-10 mM and 4000-200000 pmol/mL for styrene, 1,3-pentadiene, isoprenol and all gases, respectively (Appendix 9.4, Page 217).

4.2.6.9 Electrospray Ionisation – Mass Spectrometry

The analysis was performed by Benjamin Gleadhill (School of Chemistry, University of Nottingham). Reaction mixtures were filtered using a Vivaspin Sample Concentrator to remove enzyme and quench the reaction. Filtrates were analysed by ESI-MS operating in negative ion mode using a Varian 310-MS TQ system controlled by a Varian Workstation version 6.9. Samples were diluted in 50% water-50% acetonitrile for direct infusion at 0.05 mL/min. Ten spectra were averaged to generate the data used in the figures. The masses of the negatively charged ions discussed in the text are mevalonate 3-phosphate (227 m/z) and mevalonate (147 m/z), and the phosphate species are H_2PO_4 (97 m/z) and PO_3 (79 m/z). For tandem MS (MS-MS) acquisition, the isolation width was set at 1 m/z, and the collision energy was set at 10 eV.

4.2.6.10 Nuclear Magnetic Resonance

The analysis was performed by Shaz Aslam (School of Chemistry, University of Nottingham). The samples containing the unidentified reaction product of *P. torridus* MVD using MVA as substrate were diluted to 600 µL with 99.9% D_2O. Spectra were acquired at ambient temperature using a 500 MHz Bruker AV(III)500 spectrometer equipped with a cryoprobe. Data were processed and analysed using ACD/NMR Processor version 12.01. Authentic standards of MVA and MVA-5P, prepared at a concentration of 40 mM, were analysed in the same way.

5 Results

5.1 Phenylacrylic acid decarboxylases

5.1.1 Introduction

Two phenylacrylic acid decarboxylases (PADs), *S. cerevisiae* PAD1 and *A. niger* PadA1, raised interest for the decarboxylation of mesaconic, citraconic and itaconic acid due to similarity between subtrates known to be recognised by these enzymes and the substrates of interest. These PADs catalyse the decarboxylation of cinnamic acid and sorbic acid to styrene and 1,3-pentadiene, respectively (Figure 5.1) (Clausen *et al.*, 1994; Plumridge *et al.*, 2007; Stratford *et al.*, 2007). Sorbic acid is an aliphatic acrylic acid, similar in structure to itaconic, mesaconic and citraconic acid. Therefore these PADs may be able to catalyse their decarboxylation. To further increase the interest in these enzymes, the fact that *A. niger* showed substrate promiscuity and was able to decarboxylate several other sorbic acid derivatives (Stratford *et al.*, 2012).

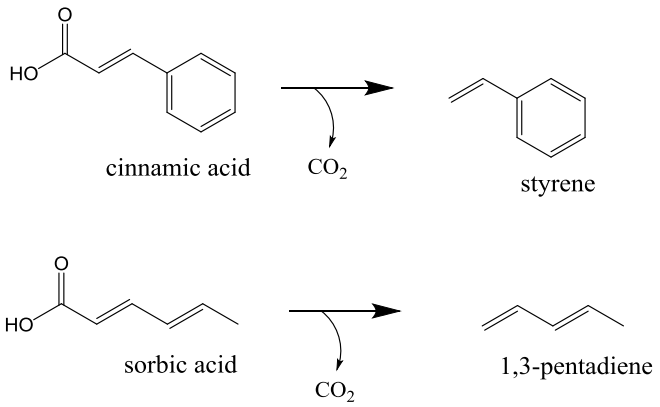


Figure 5.1: Natural activity of PADs

PADs are a recently discovered class of decarboxylases that catalyse the conversion of phenylacrylic acids to their respective alkenes *via* decarboxylation (Clausen *et al.*, 1994; Mukai *et al.*, 2010; Plumridge *et al.*, 2007). They have been found in moulds,

yeasts and bacteria and their presence is mainly correlated with resistance to weak organic acids, added as preservatives in food, given by the decarboxylation of these compounds (Clausen *et al.*, 1994; Mukai *et al.*, 2010; Plumridge *et al.*, 2007). PAD often also refers to phenolic acid decarboxylases. Although the nomenclature is confusing, the two classes of enzymes are distinct. Phenolic acid decarboxylases require the presence of a hydroxyl substituent on the phenyl moiety of the substrate, whereas PADs do not. Moreover, whereas phenolic acid decarboxylases are usually cofactor-free enzymes and the reaction mechanism is mostly known (Rodríguez *et al.*, 2010), the cofactor requirement and the reaction mechanism of PADs remains unclear[1]. Examples of applications using PADs have been reported, where the most significant is the fermentative production of unsubstituted styrene from *trans*-cinnamic acid, using engineered *E. coli* (McKenna and Nielsen, 2011).

Recent studies suggested that activity of both PAD1 and PadA1 also requires the expression of additional enzymes naturally encoded by adjacent genes. These enzymes are FDC1 and OhbA1 in *S. cerevisiae* and *A. niger*, respectively (Mukai *et al.*, 2010; Plumridge *et al.*, 2010). FDC1 refers to a putative ferulic acid decarboxylase whereas OhbA1 refers to a putative 3-octaprenyl-4-hydroxybenzoic acid decarboxylase. The interaction between the enzymes is not yet clear[1]. Initially it was believed that PAD1 and PadA1 were the main catalysts involved in sorbic and cinnamic acid decarboxylation and that FDC1 and OhbA1 were just co-catalysts (Mukai *et al.*, 2010; Plumridge *et al.*, 2007). More recent studies suggested that it may be the opposite (Leys *et al.*, 2013; McKenna and Nielsen, 2011). PADs are still poorly understood and it is still not clear if and how the pairs PAD1/FDC1 and PadA1/OhbA1 interact and whether or not they are both crucial for the reaction[1]. Therefore, in order to investigate activity of PADs on mesaconic,

[1]Just before submission of this thesis, two papers were published proposing that padA1 is not a decarboxylase *per se* but a flavin prenyl transferase that forms a prenylated flavin (prFMN). prFMN is required by OhbA1 for decarboxylation which occurs *via* 1,3 dipolar cycloaddition (Payne *et al.*, 2015; White *et al.*, 2015). This study reveals the functional relationship between the two enzymes.

citraconic and itaconic acid, both *S. cerevisiae* PAD1 and *A. niger* PadA1 were tested together with FDC1 and OhbA1, respectively.

5.1.2 Results

5.1.2.1 Enzyme expression

The genes (Appendix 9.1, Page 197-200) were amplified by PCR from genomic DNA extracted from the original organisms. For the genes encoding PAD1, FDC1 and PadA1, a standard PCR amplification was performed. Reaction conditions are indicated in Table 5.1. The amplification of the gene encoding OhbA1 required a modified protocol due to the presence of an intron that needed to be removed to enable expression in *E. coli* (Figure 5.2). Reaction conditions are indicated in Table 5.2.

Table 5.1: Reaction conditions for PCR of genes encoding PAD1, FDC1 and PadA1

	***S. cerevisiae* PAD1**	***S. cerevisiae* FDC1**	***A. niger* PadA1**
Primers	PR1 – PR2	PR3 – PR4	PR5 – PR6
[DNA]	0.06 ng/µL (*S. cerevisiae* W303)	0.24 ng/µL (*S. cerevisiae* W303)	0.24 ng/µL (*A. niger* N402)
[$MgCl_2$]	2 mM	2 mM	2 mM
[DMSO]	-	5%	-
PCR method	3 min 95°C (15 s 95°C, 30 s 58°C, 1 min 72°C) x35 5 min 72°C	3 min 95 °C (15 s 95°C, 30 s 40°C, 1.5 min 72°C) x5 (15 s 95°C, 30 s 58°C, 1.5 min 72°C) x30 5 min 72 °C	3 min 95 °C (15 s 95°C, 30 s 44°C, 1.5 min 72°C) x5 (15 s 95°C, 30 s 60°C, 1.5 min 72°C) x30 5 min 72 °C

Table 5.2: PCR conditions for gene encoding *A. niger* OhbA1

	1st PCR	**2nd PCR (Exon 1)**	**2nd PCR (Exon 2)**	**3rd PCR**
Primers	PR7 – PR8	PR7 – PR9	PR8 – PR10	PR7 – PR8
[DNA]	0.24 ng/µL (A. niger N402)	1:100 dilution (Product of 1st PCR)	1:100 dilution (Product of 1st PCR)	1:10 dilution (Exon 1: Exon2 mixture)
[$MgCl_2$]	2 mM	1.5 mM	1.5 mM	2 mM
[DMSO]	5%	-	-	-
PCR method	3 min 95°C (15 s 95°C, 30 s 50 °C, 1.5 min 72°C) x5 (15 s 95°C, 30 s 62°C, 1.5 min 72°C) x30 5 min 72°C	3 min 98°C (15 s 98°C, 2 s 60°C, 20 s 72°C) x30 5 min 72 °C	3 min 98°C (15 s 98°C, 2 s 50°C, 20 s 72°C) x5 (15 s 98°C, 2 s 62°C, 20 s 72°C) x25 5 min 72°C	3 min 98°C (15 s 98°C, 2 s 50°C, 20 s 72°C) x5 (15 s 98°C, 2 s 60°C, 20 s 72°C) x25 5 min 72°C

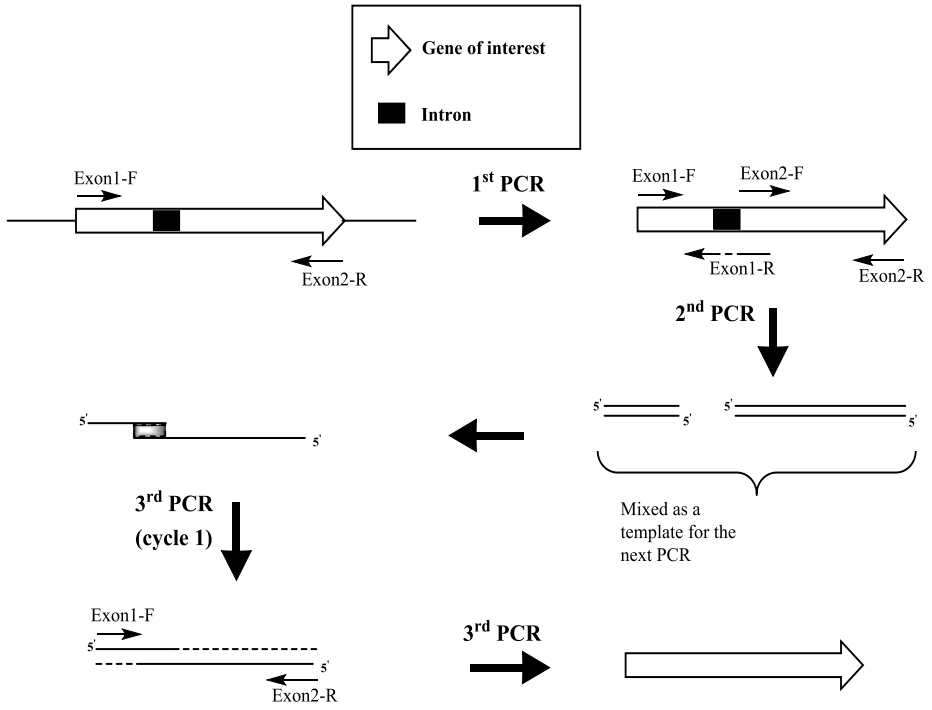


Figure 5.2: PCR protocol for intron removal
In the 1st PCR the whole gene including the intron is amplified using Exon1-F and Exon2-R primers. In the 2nd PCR, the two exons are amplified singly using respective reverse (R) and forward (F) primers. One of the primers (Exon1-R) is designed to be complementary to part of Exon1 and part of Exon2. Using this primer to amplify Exon1, the resulting fragment includes part of the Exon2 sequence and, when mixed together for another PCR reaction (3rd PCR), Exon1 and Exon2 are able to base-pair *via* the overlap. At this stage, a template of the whole gene is provided and, using Exon1-F and Exon2-R primers, the intron-free gene can be amplified.

The amplified genes were inserted into pET-20b(+) and the plasmids transformed into *E. coli*. In LB medium, the optimal conditions for *S. cerevisiae* PAD1 and FDC1 were at 37°C and 30°C, respectively, as reported previously (McKenna and Nielsen, 2011). Expression of FDC1 was very good after 3-5 h of induction, whereas a period of 5-24 h was needed for PAD1 to obtain sufficient enzyme (Appendix 9.3 - Page 207-208). A similar trend was observed for A. *niger* PadA1 and OhbA1. PadA1 required expression at 37°C for 5-24 h, and OhbA1 at 30°C for only 3-5 h (Appendix 9.3, Page 209-210). The estimated molecular weight of PAD1 and PadA1 was very similar (around 25 kDa). FDC1 and OhbA1 also had similar weights (around 55 kDa). The four enzymes were expressed at larger scale under the optimum conditions and cells were used to prepare crude extracts for enzymatic assays.

5.1.2.2 Enzyme activity

The crude extracts containing a mixture of PAD1/FDC1 or PadA1/OhbA1, were first tested for the decarboxylation of *trans*-cinnamic and sorbic acid. As expected, formation of styrene and 1,3-pentadiene, respectively, was observed only when using crude extracts containing the enzymes of interest, and no product was detected in the control samples (Figure 5.3). The conversion yield was around 40-50% for cinnamic acid and close to 100% for sorbic acid. Both pairs were then tested for MAA formation by decarboxylation of itaconic, mesaconic and citraconic acid (10 mM) under the same reaction conditions. Substrate consumption and MAA formation were monitored by HPLC. Unfortunately, formation of MAA was not observed either with *S. cerevisiae* or *A. niger* PADs, even though the detection limit for MAA was <0.1 mM. Some minor consumption of substrates (0.08% maximum) was observed but the same was observed in the controls. These results indicated that PADs do not catalyse the target reactions.

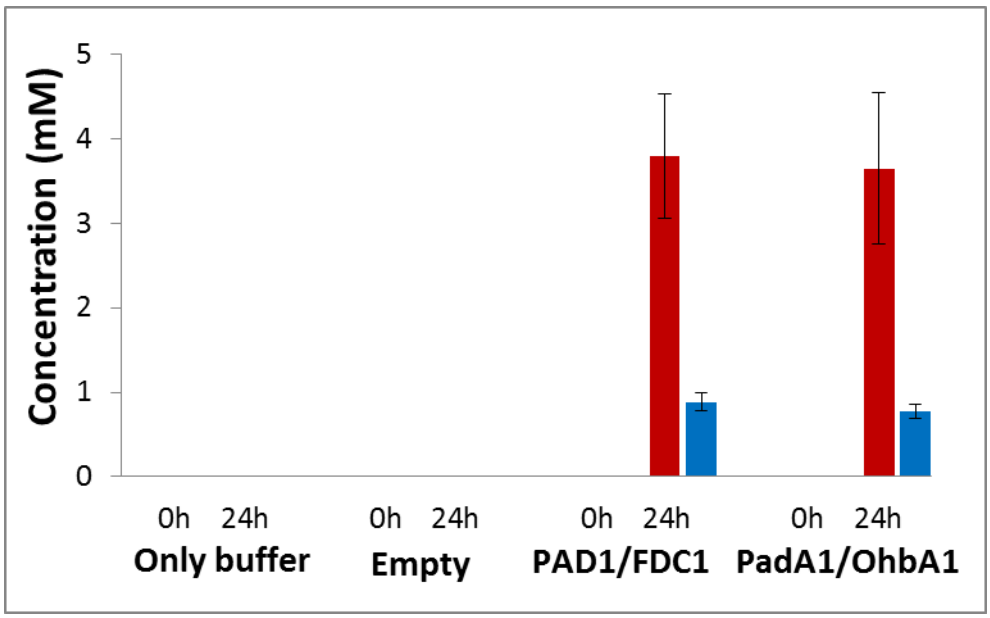


Figure 5.3: *S. cerevisiae* and *A. niger* PADs enzyme activity confirmation
S. cerevisiae PAD1/FDC1and *A. niger* PadA1/OhbA1 were assayed using *trans*-cinnamic acid (6.7 mM) and sorbic acid (1 mM) as the substrates. Respective formation of styrene (red bars) and 1,3-pentadiene (blue bars) is shown. The reaction was performed in 50 mM phosphate buffer (pH 6.5) at 30˚C for 24 h with gentle shaking. (Only buffer) refers to the assay performed in the absence of crude extract. (Empty) refers to the assay performed with crude extract from *E. coli* containing empty vector. (PAD1/FDC1) and (PadA1/OhbA1) refers to the assay performed with the mixture of crude extracts containing the enzymes of interest. For each sample, results are provided for 0 and 24 h. Values are the average of three replicates.

5.1.2.3 Activity of individual enzymes

In the interest of gaining more insights into PADs, the enzymes were tested in the absence of the respective suggested co-catalysts. Crude extracts containing the single enzymes were mixed with *trans*-cinnamic and sorbic acid, and the formation of styrene and 1,3-pentadiene, respectively, was monitored by GC-MS after 24 h (Figure 5.4).

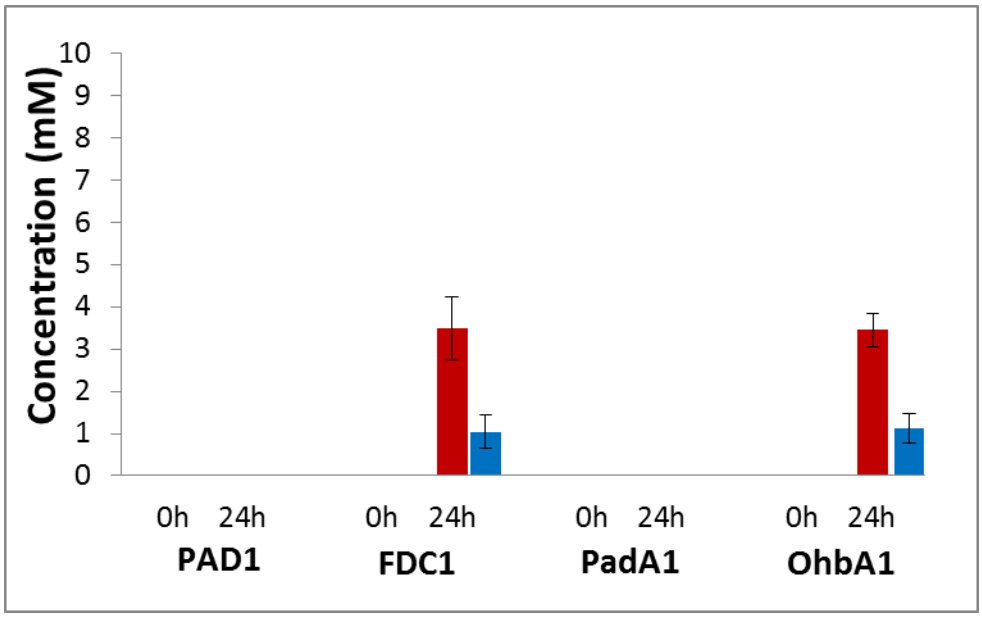


Figure 5.4: Activity of individual PADs from *S. cerevisiae* and *A. niger*
S. cerevisiae PAD1 and FDC1and *A. niger* PadA1 and OhbA1 were assayed using a *trans*-cinnamic acid (6.7 mM) and sorbic acid (1 mM) as the substrates. Respective formation of styrene (red bars) and 1,3-pentadiene (blue bars) is shown. The reaction was performed in 50 mM phosphate buffer (pH 6.5) at 30°C for 24 h with gentle shaking. (PAD1), (FDC1), (PadA1) and (ohbA1) refer to the assay performed with the crude extracts containing the enzymes of interest. For each sample, results are provided for 0 and 24 h. Values are the average of three replicates.

When PAD1 and PadA1 were tested alone, neither styrene nor 1,3-pentadiene was observed in the sample (Figure 5.4). This indicated that the enzymes are not directly needed for decarboxylation. By contrast, FDC1 and OhbA1 were able to decarboxylate both sorbic and *trans*-cinnamic acid on their own, since 1,3-pentadiene and styrene could be detected at the end of the reaction (Figure 5.4). For *S. cerevisiae* this had already been directly suggested (McKenna and Nielsen, 2011), but for *A. niger* PadA1/OhbA1, this was the first time that the enzymes were expressed and studied individually in heterologous systems. With

this experiment, it was demonstrated that, contrary to a previous report (Plumridge *et al.*, 2007), OhbA1 is the main enzyme involved in the decarboxylation of sorbic and *trans*-cinnamic acid, and not PadA1. Interestingly, conversion yields using both FDC1 and OhbA1 were 40-50% for *trans*-cinnamic acid and close to 100% for sorbic acid, as for the assays performed using mixtures of PAD1/FDC1 and PadA1/OhbA1, suggesting that that the presence of PAD1 and PadA1 does not affect activity. This does not explain why PAD1 and PadA1 are naturally co-expressed with FDC1 and OhbA1 in the original microorganisms and how they interact with the main catalyst. Since the assay was done in crude extract, the effect of PAD1 and PadA1 may have been complemented by the presence of *E. coli* UbiX, as already suggested previously (McKenna and Nielsen, 2011). UbiX is a protein that has ~50% similarity with PAD1 and PadA1 and is involved in the synthesis of ubiquinone (Rangarajan *et al.*, 2008).

5.1.2.4 Studies on purified OhbA1

In order to better understand the function of *A. niger* PADs and the possible involvement of PadA1 or other *E. coli* homologue such as UbiX, OhbA1 was purified and assayed at different conditions. This work was done in collaboration with Prof David Archer and Dr Michaela Novodvorska (University of Nottingham, School of Biosciences). Gene cloning and purification procedures to obtain functional, purified OhbA1 were performed by me, Dr Novodvorska performed all the enzymatic assays. This study is still in progress and, at the moment, only preliminary data are available.

Purified OhbA1 (Figure 5.5) was able to decarboxylate cinnamic acid. However, its activity was drastically enhanced by the addition of crude extract from non-engineered *E. coli*. When the crude extract was filtered with a 2 kDa cut-off membrane, only the flow-through fraction was able to increase activity, whereas the retained solution had no effect. This suggested that a molecule with a size lower than 2 kDa is probably required. This molecule may be a cofactor or a small peptide

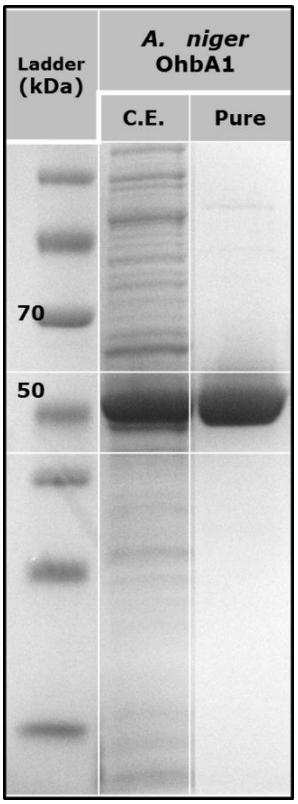


Figure 5.5: Purification of *A. niger* OhbA1.
SDS-PAGE analysis of *A. niger* OhbA1 in crude extract (C.E.), and after purification on a Ni Sepharose column (Pure). The level of migration corresponding to OhbA1 is marked between white lines.

To further investigate this, purified OhbA1 was first tested in the presence of different metals which are commonly found in decarboxylases. Mg^{2+} was found to increase activity (Figure 5.6). The enzyme was then tested with commercially available organic cofactors and thiols, in the presence of Mg^{2+} (Figure 5.7). With ATP and NADH, the enzyme retained full or partial activity. With FAD and FADH, the activity was lost. This suggests that FAD and FADH may bind at the active site and prevent binding of the substrate or of some other cofactor. In contrast, styrene formation was enhanced drastically in the presence of thiols such as dithiothreitol, 2-mercaptoethanol, homocysteine and glutathione (Figure 5.7). This suggested that the activity of the enzyme could be dependent on the redox state of cysteines perhaps around the active site and/or the binding pocket.

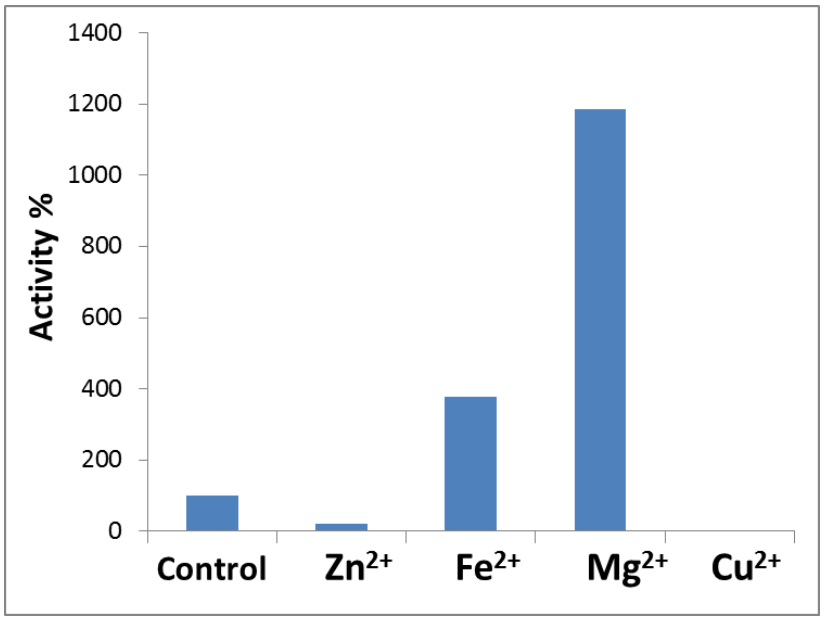


Figure 5.6: Effect of metals on OhbA1 activity
Purified *A. niger* OhbA1 was assayed using cinnamic acid (1 mM) in the presence of divalent cation metals (1 mM). The reaction was performed in 50 mM phosphate buffer (pH 6.5) at 28°C for 24 h with gentle shaking. Formation of styrene was then monitored by GC-MS. The blue bars indicate the percentage of activity. 100% activity refers to the amount of styrene detected when the assay was performed with no metals (0.015 mM) (Control). The experiments are preliminary and no replicates have been analysed.

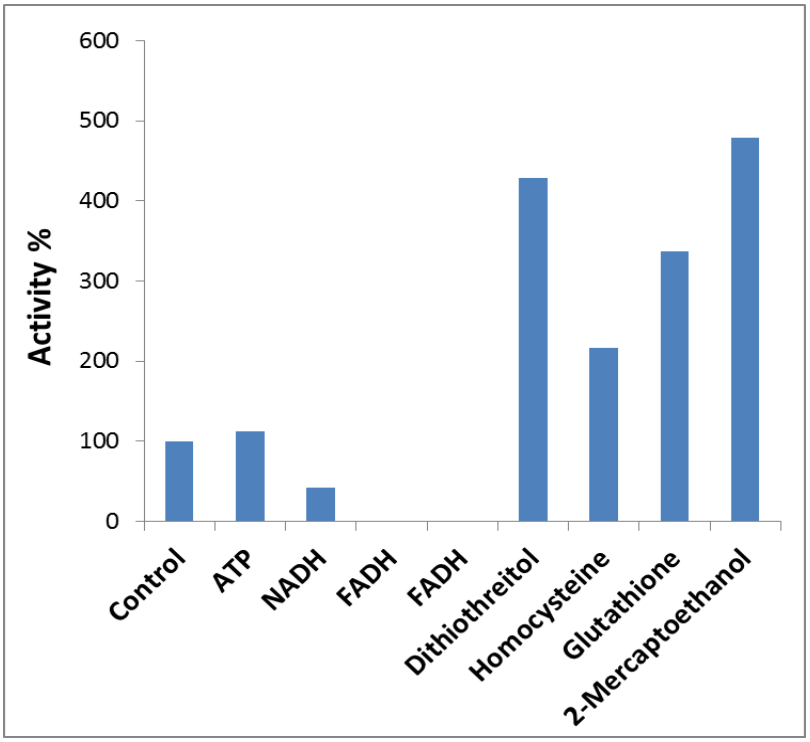


Figure 5.7: Effect of organic cofactors and thiols on OhbA1 activity
Purified *A. niger* OhbA1 was assayed using cinnamic acid (1 mM) in the presence of organic cofactors (5-10 µM) and thiols (1 mM). Mg^{2+} (1 mM) was added in all samples. The reaction was performed in 50 mM phosphate buffer (pH 6.5) at 30°C for 24 h with gentle shaking. Formation of styrene was then monitored by GC-MS. The blue bars indicate the percentage of activity. 100% activity refers to the amount of styrene detected when the assay was performed with no organic cofactors or thiols (0.18 mM) (Control). The experiments are preliminary and no replicates have been analysed.

5.1.2.5 Structural analysis

A homology model of OhbA1 was obtained using I-TASSER (Roy *et al.*, 2010; Zhang, 2008). Cinnamic acid was docked *in silico* to predict the position of the active site and to identify cysteines that could be involved in the enzyme function. Interestingly, Cys316 is very close to cinnamic acid (Figure 5.8). Therefore, mutagenesis of Cys316 to Ala using the QuickChange Lightening Site-Directed Mutagenesis Kit, which is based on the principle of whole plasmid amplification with mutagenic primers (Papworth *et al.*, 1996), was performed. The Cys316Ala mutant was then tested on *trans*-cinnamic acid in the presence of Mg^{2+}, and resulted in an 80% reduction of activity (0.036 mM styrene formed after 24 h against 0.28 mM with the wild type). The homology model also revealed the presence of a well-structured Mg^{2+} binding domain, consisting of Glu233, Asn168 and His191 (Figure 5.8), and a big cleft positioned on top of the predicted active site.

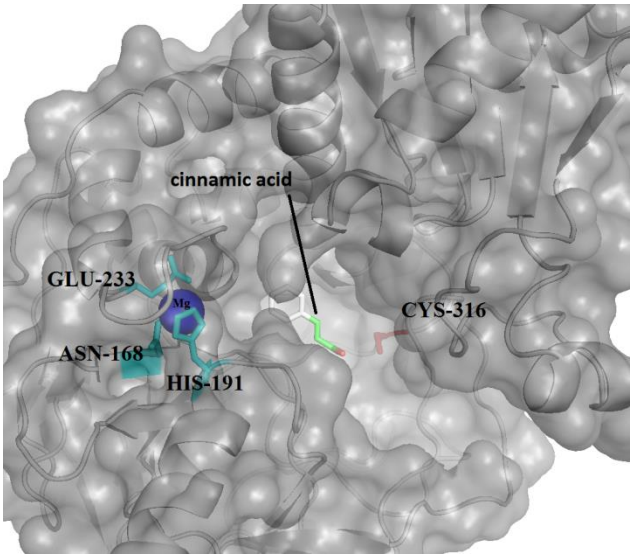


Figure 5.8: Homology model of OhbA1
The picture focuses on the predicted catalytic cleft. Important residues are highlighted: cyan for residues involved in the binding of Mg^{2+}, and red for Cys316, potentially involved in catalysis. Cinnamic acid was docked into the active site.

5.1.3 Discussion

The phenyl acrylic decarboxylases (PADs) from *S. cerevisiae* and *A. niger*, PAD1/FDC1 and PadA1/OhbA1, were chosen as candidates

to catalyse the conversion of itaconic, mesaconic and citraconic acid to MAA, as they are active on similar aliphatic acrylic acids. The enzymes were successfully expressed in soluble form in *E. coli* without the need of extensive optimisation, and were subsequently tested on their natural substrates to confirm their expression as active catalysts. All enzymes were finally assayed on itaconic, mesaconic and citraconic acid for the production of MAA but, unfortunately, no activity was observed.

Recent studies on PADs have suggested a few important structural features required for substrates to be successfully decarboxylated (Stratford *et al.*, 2012). For example, beyond a hydrophilic carboxylic acid head, the remaining part of the substrates should be mostly hydrophobic. Itaconic, mesaconic and citraconic acid do not exhibit these structural features and present a polar and charged carboxylic group. Moreover, the alkenyl bond between C2 and C3 should be in *trans* configuration (Stratford *et al.*, 2012). The alkenyl bond of both itaconic and citraconic acid is in *cis* configuration, instead. Finally, substrate length between ~6 and 9 Å were observed to increase enzyme activity, but the tested substrates were smaller. Although this study was performed on *A. niger* PadA1/OhbA1 system (Stratford *et al.*, 2012), it is very likely that *S. cerevisiae* PAD1 and FDC1 require the same features for their substrates, since they are very similar enzymes. This may explain why none among the tested PADs was active on the substrates of interest.

In order to gain more insights into this poorly characterised sub-class of decarboxylases, further investigation on the enzymes was performed. It was demonstrated that FDC1 and OhbA1 are the main enzymes involved in the decarboxylation of cinnamic and sorbic acid, and that the respective PAD1 and PadA1 have no or indirect role. This had already been suggested for *S. cerevisiae* (McKenna and Nielsen, 2011) but it is a new discovery for the *A. niger* PAD system, where PadA1 was so far believed to be the decarboxylating enzyme (Plumridge *et al.*, 2007).

The activity of purified *A. niger* OhbA1 was also investigated. The presence of crude extract from non-engineered *E. coli* increased enzymatic activity. However, after filtering the extract with a 2 kDa cut-off membrane, only the flow-through fraction was able to increase activity. Therefore, it was not a protein-protein interaction to enhance OhbA1 activity but rather the presence of a smaller cofactor. Further investigation revealed that the activity of the enzyme is enhanced by the presence of divalent cations, Mg^{2+} above all, and by the presence of thiols (Figure 5.6, Page 67 and Figure 5.7, Page 67). Other cofactors such as ATP, NADH, FAD and FADH, found in many decarboxylation reactions, have no or negative effect on the enzyme. Since with FAD and FADH, the activity was lost, it is possible that these cofactors bind into the active site and prevent the binding of the substrate or other cofactors. This information is crucial to gain more insights into OhbA1 function but still does not explain how PadA1 can be involved in the reaction and why it is co-expressed with OhbA1 in *A. niger*[1].

Very recently, a study on *S. cerevisiae* PAD1/FDC1 system suggested that the functional role of PAD1 is to produce a cofactor for FDC1, required for decarboxylation (Lin *et al.*, 2015). The authors also suggest that this cofactor is a novel FMN derivative, but they were not able to characterise it fully. Since the *S. cerevisiae* PAD system is very similar to the *A. niger* system, it is possible that a similar mechanism occurs for PadA1 and OhbA1 as well[1]. Analysis of the homology model of OhbA1 showed the presence of a metal binding domain into the active site, consisting of Glu233, Asn168 and His191 (Figure 5.8). This explains why Mg^{2+} enhanced activity of OhbA1. However, the domain is positioned quite far from where cinnamic acid binds into the predicted binding site. Therefore, its presence, rather than a direct role in catalysis, may have either a structural role or may be involved in the coordination of a cofactor.

[1] Just before submission of this thesis, two papers were published proposing that padA1 is not a decarboxylase *per se* but a flavin prenyl transferase that forms a prenylated flavin (prFMN). prFMN is required by OhbA1 for decarboxylation which occurs *via* 1,3 dipolar cycloaddition (Payne *et al.*, 2015; White *et al.*, 2015). This study reveals the functional relationship between the two enzymes

Moreover, OhbA1 catalytic cleft is big (Figure 5.8), and it is possible that it could accommodate a FMN-like cofactor, as suggested in FDC1 (Lin *et al.*, 2015). Finally, the homology model of OhbA1 also allowed identification of a residue, Cys316, that could potentially be involved in catalysis. Alanine mutagenesis of this residue resulted in an 80% decrease in enzyme activity. This would explain why the presence of thiols, able to affect the redox state of Cys316, can enhance activity of OhbA1.

PADs are an unusual, still not weel characterised, class of enzymes which catalyses the decarboxylation of aromatic and aliphatic acrylic compounds without the need of ring activation. This was very attractive for the decarboxylation of itaconic, mesaconic and citraconic acid. Although this study has shown that, at present, this route does not seem feasible, several information regarding potential important residues and potential cofactor requirements of PADs was unveiled. This may provide, in future, more insights into their use for the bioproduction of MAA and into how to improve processes were PADs are already employed, such as the bioproduction of styrene and other monounsaturated alkenes(Leys *et al.*, 2013; McKenna and Nielsen, 2011; McKenna *et al.*, 2014). Moreover, PADs could potentially be employed to produce other useful chemicals with a chemical conformation similar to that of the natural substrates. An interesting example would be the conversion of 2,4-pentadienoic acid to butadiene (Figure 5.9), an important monomer used in the production of synthetic rubber. This is an exciting prospect, since biocatalytic routes for the production of 2,4-pentadienoic acid have been already described (Osterhout *et al.*, 2013)

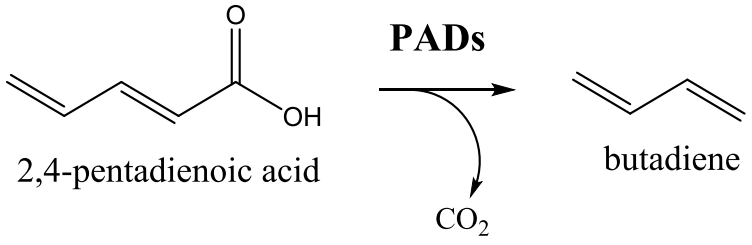


Figure 5.9: Potential formation of butadiene catalysed by PADs

5.2 *cis*-Aconitate decarboxylase

5.2.1 Introduction

cis-Aconitate decarboxylase (CAD) (EC 4.1.1.6) was selected for the decarboxylation of mesaconic and citraconic acid. What raised interest on this enzyme was the substrate similarity between its natural substrate, *cis*-aconitic acid, and the two substrates of interest (Figure 5.10). *cis*-Aconitic acid presents the same dicarboxylic/acrylic moiety as the two target substrates, although it also carries an extra carboxylic acid on C5. It must be noted that, whereas citraconic acid presents the same *cis* configuration of the dicarboxylic moiety, mesaconic acid presents a *trans* configuration. Nevertheless, CAD was tested on both substrates.

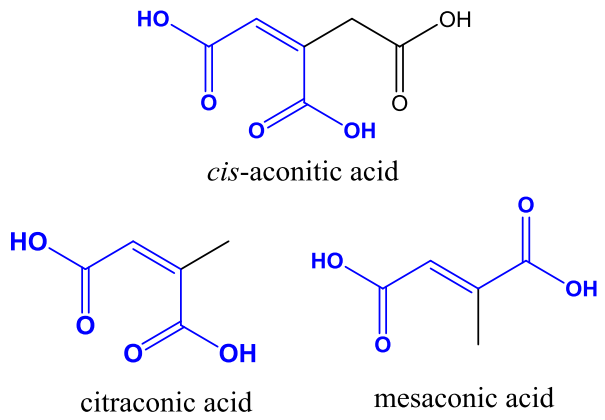


Figure 5.10: Structure similarity of citraconic and mesaconic acid with *cis*-aconitic acid

CAD is a crucial catalyst for the production of itaconic acid (Bentley and Thiessen, 1957; Blumhoff *et al.*, 2013; Chun and Kim, 2014; Huang *et al.*, 2014; Li *et al.*, 2013a; Li *et al.*, 2012a; Steiger *et al.*, 2013; van der Straat *et al.*, 2014; Vuoristo *et al.*, 2015). The enzyme naturally catalyses the decarboxylation of the TCA cycle intermediate, *cis*-aconitic acid, to itaconic acid (Figure 5.11) in what it is believed to be a cofactor-free reaction (Bentley and Thiessen, 1957).

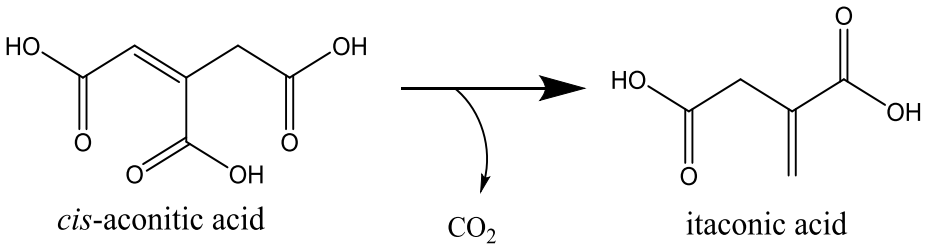


Figure 5.11: Natural activity of CAD

CAD activity was discovered in *Aspergillus terreus*, when a crude extract preparation was found to convert *cis*-aconitic acid to itaconic acid (Bentley and Thiessen, 1957; Bonnarme *et al.*, 1995; Jaklitsch *et al.*, 1991). More recently, *A. terreus* CAD (CadA1) was isolated, purified and overexpressed in *S. cerevisiae*, confirming its role in itaconic acid production (Dwiarti *et al.*, 2002; Kanamasa *et al.*, 2008). The gene sequence, optimal conditions for activity and information regarding a few inhibitors (mostly heavy metals) are now available and the enzyme has been classified as a member of the MmgE/PrpD family, which is still poorly characterized (Lohkamp *et al.*, 2006; Okabe *et al.*, 2009).

During attempts of improving itaconic acid production yields, CadA1 was heterologously expressed in different systems (Li *et al.*, 2012a; Steiger *et al.*, 2013), including *E. coli* (Vuoristo *et al.*, 2015). However, a comprehensive study revealing its catalytic mechanism, structure and activity is not yet available. The cause of this missing information could reside in the fact that the enzyme is very unstable (Steiger *et al.*, 2013) and its expression in non-natural hosts requires significant optimization (van der Straat *et al.*, 2014; Vuoristo *et al.*, 2015). Purification of CadA1 is, therefore, difficult, and structural and mechanistic studies very hard to perform.

5.2.2 Results

5.2.2.1 Enzyme expression

The gene encoding for CadA1 contains an intron (Appendix 9.1, Page 201). Therefore, it was amplified using a protocol similar to

that used for *A. niger* OhbA1 (Figure 5.2, Page 62). Reaction conditions and primers are indicated below (Table 5.3).

Table 5.3: PCR conditions for the gene encoding CadA1

	1st PCR	**2nd PCR (Exon 1)**	**2nd PCR (Exon 2)**	**3rd PCR**
Primers	PR11 – PR12	PR11 – PR13	PR12 – PR14	PR11 – PR12
[DNA]	0.24 ng/µL (*A. terreus*)	1:1000 dilution (Product of 1st PCR)	1:10 dilution (Product of 1st PCR)	1:10 dilution (Exon 1: Exon2 mixture)
[$MgCl_2$]	2 mM	2 mM	1.5 mM	1.5 mM
[DMSO]	5%	-	-	-
PCR method	3 min 95°C (15 s 95°C, 30 s 50°C, 1.5 min 72°C) x5 (15 s 95°C, 30 s 61°C, 1.5 min 72°C) x30 5 min 72°C	3 min 98°C (15 s 98°C, 2 s 60 °C, 20 s 72°C) x30 5 min 72 °C	3 min 98°C (15 s 98°C, 2 s 50°C, 20 s 72°C) x5 (15 s 98°C, 2 s 61°C, 20 s 72°C) x25 5 min 72°C	3 min 98°C (15 s 98°C, 2 s 50°C, 20 s 72°C) x5 (15 s 98°C, 2 s 60°C, 20 s 72°C) x25 5 min 72°C

After insertion of the gene into pET-20b(+) and transformation of *E. coli*, overexpression of CadA1 was problematic as expected (Steiger *et al.*, 2013). Expression in LB medium at 30 and 37°C resulted in insoluble CadA1, even after brief induction periods (Appendix 9.3, Page 212). However, lowering the temperature to 20°C allowed a significant enhancement in solubility without the need to test different media or promoters. Although most of the enzyme was still found as insoluble protein, a stable and consistent production of soluble catalyst was observed from the earliest stages of induction (Appendix 9.3, Page 212). The molecular weight of CadA1 was around 55 kDa, consistent with previous reports (Dwiarti *et al.*, 2002; Kanamasa *et al.*, 2008). Therefore, crude extracts were tested in the enzymatic assays.

5.2.2.2 Enzyme activity

CadA1 was tested for decarboxylation of *cis*-aconitic acid (Dwiarti *et al.*, 2002), by monitoring consumption of substrate and formation of itaconic acid by HPLC. Consumption of *cis*-aconitic acid was observed for both samples containing CadA1 and for the controls (Figure 5.12). This might be due to spontaneous decomposition of

the acid, since decarboxylation was observed in the controls containing only buffer. The consumption was increased in extracts of cells containing empty vector, probably due to enzymes of the TCA cycle. However, itaconic production was observed only when using CadA1, confirming that it was expressed in active form. Full conversion was reached in less than 5 h. CadA1 was then tested for decarboxylation of mesaconic and citraconic acid. Unfortunately, the enzyme showed no activity. Neither mesaconic nor citraconic acid were consumed, and no MAA formation was observed.

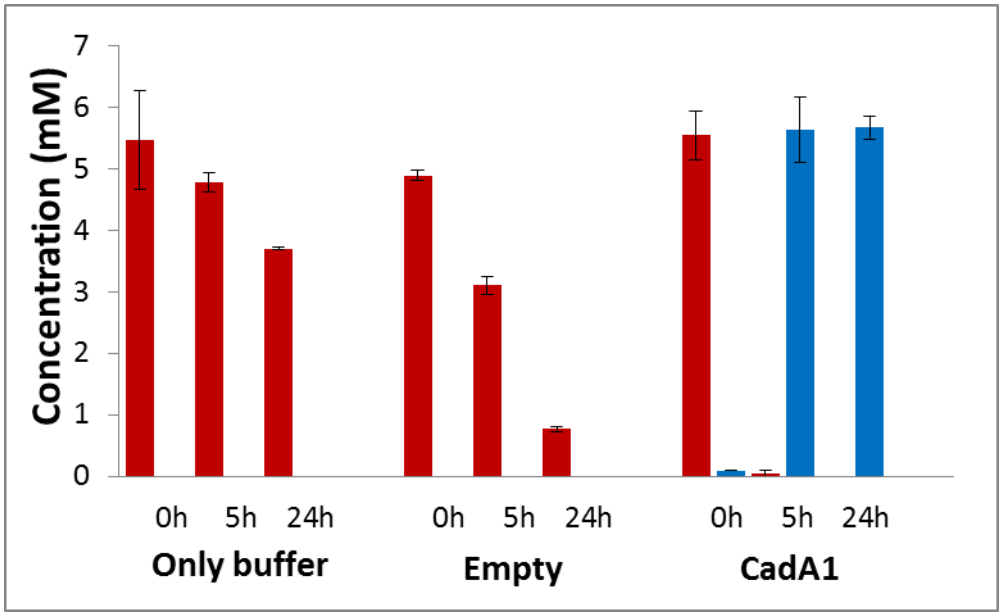


Figure 5.12: *A. terreus* CAD1 enzyme activity confirmation
A. terreus CadA1 was assayed using *cis*-aconitic acid (5 mM) as substrate. Consumption of substrate (red bars) and formation of the product itaconic acid (blue bars) are shown. The reaction was performed in 0.2 M phosphate buffer (pH 6.2) at 45°C for 24 h with gentle shaking. (Only buffer) refers to the assay performed in the absence of crude extract. (Empty) refers to the assay performed with crude extract from *E. coli* containing empty vector. (CadA1) refers to the assay performed with crude extract containing the enzyme of interest. For each sample, results are provided for 0 and 24 h. Values are the average of three replicates

Enzymology studies aimed to better understand the involvement of residues and the enzymatic mechanism of CadA1 were initially planned in collaboration with Ingenza Ltd (Edinburgh). A 3DM database containing the CadA1 sequence aligned with other 1482 protein homologues was obtained through Bio-Prodict (Netherlands) (Kuipers *et al.*, 2010). The 3DM database allows to collect and integrate information regarding the proteins included in the alignment and facilitates the exploration of sequence-structure

function relations (Jochens and Bornscheuer, 2010). However, since the results suggested that CadA1 is not a good candidate for formation of MAA, the experiments were suspended.

5.2.3 Discussion

A. terreus cis-aconitate decarboxylase (CadA1) was tested for the decarboxylation of mesaconic and citraconic acid, due to the similarity of these substrates with the natural substrate of the enzyme, *cis*-aconitic acid (Figure 5.10). Growth and expression temperature had to be lowered to 20°C to obtain a sufficient amount of soluble protein. At the time the expression was tested, it was the first time that someone attempted to use *E. coli*, but issues with stability of the enzyme had already been reported (Dwiarti *et al.*, 2002; Kanamasa *et al.*, 2008). More recently, others reported the expression of CadA1 into *E. coli* but, as observed in this study, soluble expression was hard to obtain (Okamoto *et al.*, 2014; Vuoristo *et al.*, 2015). Expression of CadA1 as active catalyst was confirmed by testing its activity on *cis*-aconitic acid. However, CadA1 was not able to decarboxylate either mesaconic or citraconic acid, as consumption of substrate or formation of MAA was not observed.

Although CadA1 has been extensively used in studies regarding itaconic acid production (Bentley and Thiessen, 1957; Blumhoff *et al.*, 2013; Chun and Kim, 2014; Huang *et al.*, 2014; Li *et al.*, 2013a; Li *et al.*, 2012a; Steiger *et al.*, 2013; van der Straat *et al.*, 2014; Vuoristo *et al.*, 2015), the enzymatic mechanism is still not known. Nevertheless, it is important to mention that an old report suggested that the decarboxylation of *cis*-aconitic acid occurs at C5, rather than the C1 (Bentley and Thiessen, 1957). Therefore, mesaconic acid and citraconic acid are not likely to be substrate of such decarboxylation because they lack of a carboxylate on C5. Since that information was dated and no experimental confirmation has been reported since, this has not detracted from the intention to test the enzyme. However, the results further strengthen what

proposed by Bentley and Thiessen, since mesaconic acid and citraconic acid were not decarboxylated.

Although the results suggested that CadA1 is not a good candidate for the formation of MAA, further investigation on this enzyme provides great interest in regards to itaconic acid production, where CadA1 is the main enzyme involved (Okabe *et al.*, 2009; Steiger *et al.*, 2013). Therefore, there is still scope for future research, and the customised 3DM database generated in the course of this work can be used for the purpose.

5.3 *α*-Amino-*β*-carboxymuconate-*ε*-semialdehyde decarboxylase

5.3.1 Introduction

α-Amino-β-carboxymuconate-ε-semialdehyde decarboxylase (ACMSD; EC 4.1.1.45) catalyses the decarboxylation of 2-amino-3-(3-oxoprop-1-en-1-yl)-but-2-enedioate (ACMS), to 2-aminomuconate-6-semialdehyde (AMS) (Figure 5.13). This enzyme was discovered in mammals, where it controls the fate of tryptophan metabolites in the kynurenine pathway for the production of NAD (Nishizuka *et al.*, 1970). ACMS is an intermediate of this pathway and can undergo a slow spontaneous conversion to quinolinic acid (QA), the universal precursor for the *de novo* synthesis of NAD. The enzyme ACMSD competes with this reaction by converting ACMS to AMS (Figure 5.13).

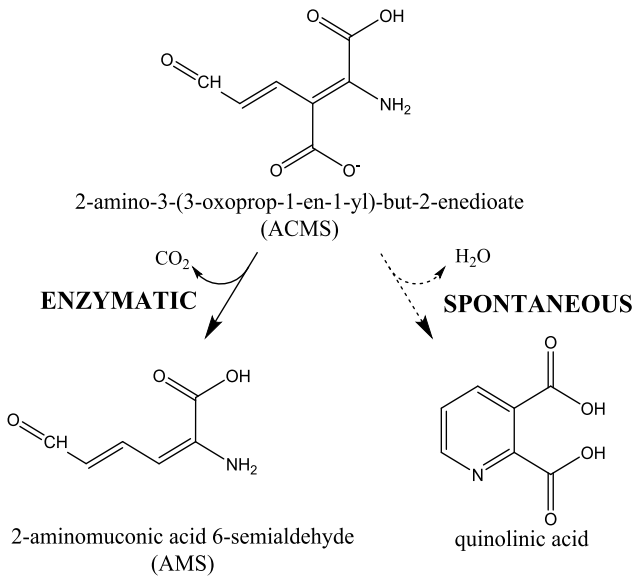


Figure 5.13: Natural activity of ACMSD

When ACMSD is fully active, it directs the metabolic flux away from QA to energy production, where AMS is converted to acetyl-CoA, which will then enter the TCA cycle (Mehler *et al.*, 1964). If activity of ACMSD is reduced, an abnormal QA concentration is

observed. In mammals, a high concentration of QA in body fluids can be linked to numerous diseases such as diabetes (Mehler *et al.*, 1964) and Alzheimer's (Guillemin *et al.*, 2003).

More recently, ACMSD was also discovered in bacteria, where it was found to have an important role in the 2-nitrobenzoic acid pathway for the biodegradation of nitroaromatic compounds (Iwaki and Hasegawa, 2007; Muraki *et al.*, 2003). Due to the importance of ACMSD in the mammalian and prokaryotic pathways, the enzyme has been extensively studied and a high level of information regarding structure, inhibition and catalytic mechanism is known (Fukuwatari *et al.*, 2004a; Fukuwatari *et al.*, 2004b; Garavaglia *et al.*, 2009; Huo *et al.*, 2013; Huo *et al.*, 2012; Iwaki and Hasegawa, 2007; Li *et al.*, 2007; Li *et al.*, 2005; Martynowski *et al.*, 2006; Muraki *et al.*, 2003; Pucci *et al.*, 2007; Tanabe *et al.*, 2002). However, instability of both substrate and product of the enzyme precluded the possibility to study ligand-bound crystal structures and some structural information was only inferred. The enzyme was classified as the first enzyme known to perform an oxygen-independent, non-oxidative decarboxylation with divalent metal cofactors (zinc, cobalt, and iron) (Li *et al.*, 2005), and as a member of a new protein subfamily in the amidohydrolase superfamily. This represents a novel non-hydrolytic C−C bond breaking activity (Li *et al.*, 2006; Liu and Zhang, 2006; Martynowski *et al.*, 2006) where no organic cofactor is needed.

The three-dimensional structure of ACMSD is nearly identical in both humans (Garavaglia *et al.*, 2009) and in *Pseudomonas fluorescens* (Huo *et al.*, 2013; Huo *et al.*, 2012; Martynowski *et al.*, 2006), consisting of an $(\alpha/\beta)^8$-barrel core with a bound zinc ion coordinated by three His, one Asp and one water molecule, all involved in catalysis (Huo *et al.*, 2012). Additionally, thanks to a small mobile insertion domain, an Arg intrudes into the active site and interacts with one of the negatively charged carboxylates of ACMS (Martynowski *et al.*, 2006). ACMSD functions as a dimer and the second carboxylate of ACMS is instead coordinated by an Arg

from the neighbouring subunit (Huo *et al.*, 2013; Huo *et al.*, 2014). The two Arg are critical for proper substrate binding at the enzyme active site and to prevent its autocyclisation to QA. Other residues have been suggested to be important in recognising ACMS (Li *et al.*, 2006) and to function as a gate to allow ligand admission (Martynowski *et al.*, 2006).

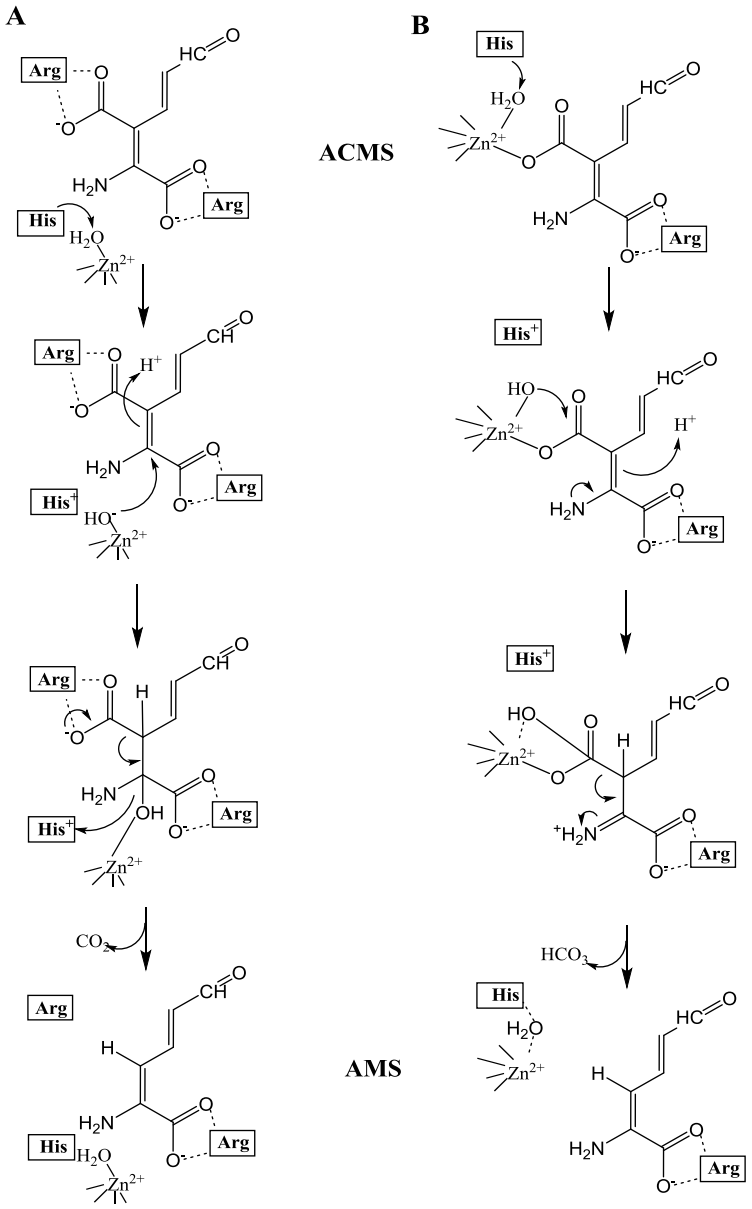


Figure 5.14: Suggested enzymatic mechanisms of ACMSD

Two catalytic mechanisms were initially suggested (Figure 5.14) (Martynowski *et al.*, 2006), but more recent studies seem to confirm the mechanism shown in Figure 5.14A (Huo *et al.*, 2013). The metal

bound water molecule is deprotonated to the hydroxide ion with the assistance of one of the previously mentioned His, which acts as a general acid/base catalyst. The hydroxide ion performs a nucleophilic attack on the C2-C3 double bond of ACMS with concomitant protonation at C3, aided by Asp. Essentially, a water molecule is added across the double bond with the hydroxyl group at C2 and the new proton at C3, generating a substrate-based tetrahedral intermediate. Collapse of the tetrahedral intermediate initiates the decarboxylation that produces AMS, and regenerates the metal centre. In this mechanism, the Arg residues present on the mobile insertion domain interacts with the C3 carboxylate of the substrate, released as CO_2.

In the second proposed mechanism (Figure 5.14B), the C3 leaving carboxylate binds next to the zinc ion with one of its oxygen atoms directly ligated to the metal. The water ligand is activated by a His residue to become a more nucleophilic hydroxide ion, which attacks the carbon atom of the leaving carboxylate group. With the concomitant protonation of C3, a tetrahedral intermediate, different from that mentioned above, is generated. The HCO_3 group of the intermediate is stabilized by the direct coordination to the Zn ion. The subsequent exchange of zinc-bound HCO_3 with a water molecule facilitates the dissociation of the product and restores the water ligand. In this route, the C3 carboxylate is released in the form of HCO_3 and the Arg residue present on the mobile insertion domain interacts with the C2 carboxylate group. Compared to the preferred mechanism (Figure 5.14A), this mechanism might be energetically unfavourable, because it is difficult to believe that a negatively charged hydroxyl group attacks a negatively charged carboxylate.

What raised interest in ACMSD is that, if the suggested mechanism (Figure 5.14A) is true, the enzyme could be able to decarboxylate mesaconic or citraconic acid. These two compounds present a similar dicarboxylic moiety to that of ACMS (Figure 5.15). The difference is that they lack an amino substituent, where they have a methyl group instead, and an aldehyde moiety. The

preferred mechanism does not indicate that the amino and aldehyde moieties are directly involved in the reaction (Figure 5.14A), and this is promising for the potential decarboxylation of mesaconic and citraconic acid. However, both moieties may still be indirectly involved in the mechanism. Both could delocalise electrons away from the double bond favouring the nucleophilic attack by water. The amino moiety may favour decarboxylation by destabilizing the tetrahedral intermediate formed during the reaction (Figure 5.14A). Moreover, both moieties may be involved in substrate binding and enzyme specificity (Li *et al.*, 2006). For all these reasons, decarboxylation of mesaconic and citraconic acid may be characterised by a high K_m and lowe K_{cat} but, since at this tage, any basal activity for MAA formation would be enough, this did not detract from the interest for the enzyme.

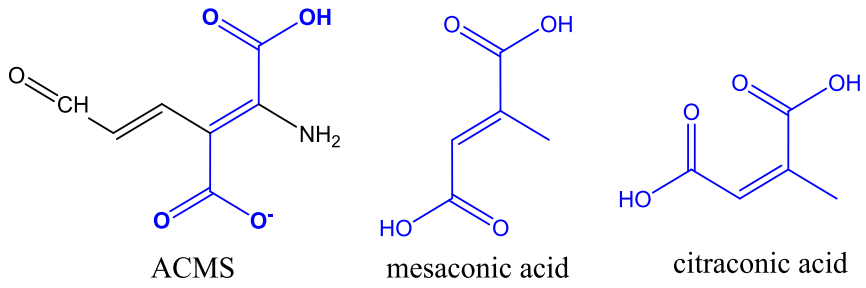


Figure 5.15: Structure similarity of mesaconic and citraconic acid with ACMS

As further confirmation that this enzyme may be a good candidate for mesaconic or citraconic decarboxylation, several inhibitors are able to bind into the active site and also present a similar dicarboxylic moiety (Fukuwatari *et al.*, 2004a; Fukuwatari *et al.*, 2004b; Fukuwatari *et al.*, 2002; Huo *et al.*, 2013; Shibata *et al.*, 2007). These compounds (Figure 5.16) are characterised by a phenyl moiety that cannot receive water addition and, therefore, they act as inhibitors. In the same way, mesaconic and citraconic acid could be able to bind the enzyme but, differently to the inhibitors, could potentially be attacked by the hydroxide ion. Interestingly, most of the known inhibitors of ACMSD present a *cis*

configuration of the two carboxylates. Therefore, citraconic acid (*cis*) may bind to ACMSD better than mesaconic acid (*trans*).

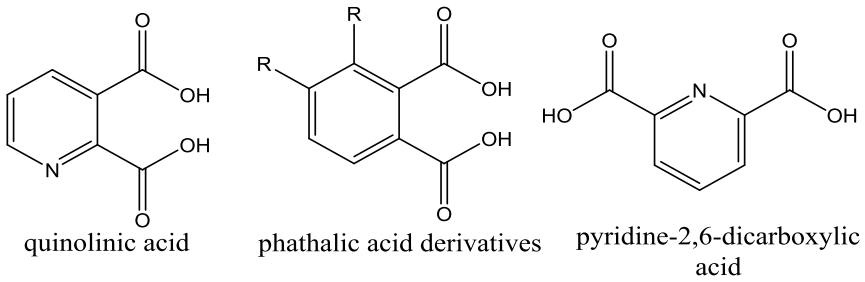


Figure 5.16: Inhibitors of ACMSD

5.3.2 Results

5.3.2.1 Enzyme expression

P. fluorescens ACMSD is better characterised than the human counterpart and has already been successfully expressed in *E.coli* (Huo *et al.*, 2013; Huo *et al.*, 2012; Martynowski *et al.*, 2006). For these reasons, it was chosen for the experiments. ACMS, the natural substrate of ACMSD, is very unstable and cannot be purchased. Therefore, to test the natural activity of ACMSD, ACMS was generated using an additional enzyme, 3,4-hydroxyanthranilinic acid dioxygenase (HAO), that converts 3-hydroxyanthranilic acid (3-HAA) to ACMS (Muraki *et al.*, 2003) (Figure 5.17). Both genes were synthesised (Appendix 9.1– Page 202-203), inserted into pET-20b(+) and transformed into *E. coli*. Expression of the enzymes was performed at 28°C, according to previously reported conditions (Huo *et al.*, 2013; Muraki *et al.*, 2003), even though the *E. coli* strain and the plasmid were different. The production of a significant amount of soluble ACMSD was observed even at early stages of induction, reaching the highest concentration after 5 h (Appendix 9.3, Page 213). However, HAO expression was not as expected, and no soluble protein was observed at 28°C (Appendix 9.3, Page 214). In order to achieve an acceptable level of soluble HAO, a temperature of 18°C for both growth and enzyme induction was

needed (Appendix 9.3, Page 214). The enzyme was still mostly insoluble, but the soluble fraction contained enough active enzyme to generate ACMS for the enzymatic assays. The molecular weights of ACMSD and HAO were about 38 and 24 kDa, respectively, consistent with previous reports (Muraki *et al.*, 2003).

5.3.2.2 Enzyme activity

ACMSD and HAO were expressed under the optimised conditions to prepare crude extracts. HAO crude extract was used to produce ACMS from 3-HAA (Figure 5.17). When ACMS production was completed, ACMSD crude extract was added and ACMS consumption monitored at 360 nm. ACMS production was only observed when HAO crude extract was present and ACMS was only consumed when ACMSD crude was added (Figure 5.17), at a rate of 0.44 nmol s^{-1} $mg_{(total\ protein)}^{-1}$. This confirmed that the enzyme was expressed in active form. Therefore, ACMSD was tested on mesaconic and citraconic acid (10 mM). Substrate consumption and product formation were monitored by HPLC. Unfortunately, neither MAA formation nor substrate consumption was observed. Since the activity of ACMSD can be increased by adding zinc (Li *et al.*, 2005; Martynowski *et al.*, 2006), the same enzyme assay was performed after incubation of the catalyst with $ZnCl_2$. However, this did not make any difference, since MAA was not formed.

It must be mentioned that, following the suggested ACMSD ligand binding system and enzymatic mechanism (Huo *et al.*, 2013; Martynowski *et al.*, 2006), it is possible that the product formed upon mesaconic or citraconic acid decarboxylation is not MAA but crotonic and isocrotonic acid, respectively. In fact, since mesaconic and citraconic acid could potentially bind in two different orientations into ACMSD's active site, the subsequent decarboxylation would occur on a different carboxylate. Only one decarboxylation would produce MAA. However, HPLC chromatograms did not show any anomalous peak that could

suggest crotonic or isocrotonic acid formation either (standard of both compounds were detectable even at 0.1 mM).

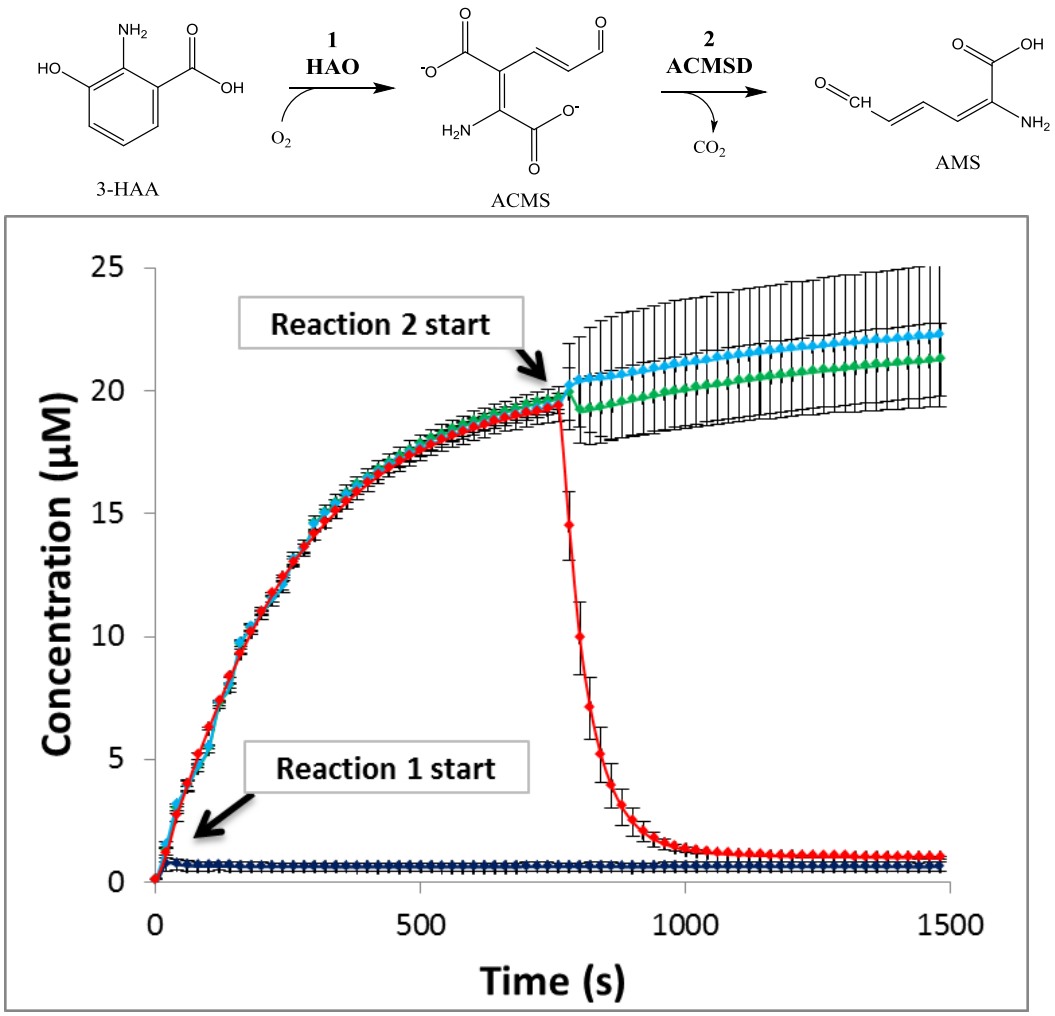


Figure 5.17: *P. fluorescens* ACMSD enzyme activity confirmation
P. fluorescens ACMSD was assayed on *α*-amino-*β*-carboxymuconate-*ε*-semialdehyde (ACMS) as substrate. ACMS was first produced from 3-hydroxyanthranilic acid (3-HAA) using 3,4-hydroanthranilic acid dioxygenase (HAO) crude extract (Reaction 1). Conversion of ACMS to 2-aminomuconate-6-semialdehyde (AMS) was then catalysed by ACMS decarboxylase (ACMSD) (Reaction 2). The reaction can be monitored at 360 nm. 3-HAA concentration was 20 µM. The reaction was performed in 25 mM HEPES buffer (pH 7) with 5% glycerol at 25°C. The blue line refers to the assay performed without addition of HAO crude extract. The green line refers to the assay performed with HAO crude extract for reaction 1 and with only water addition for reaction 2. The light blue line refers to the assay performed with HAO crude extract for reaction 1 and crude extract of *E. coli* containing empty vector for reaction 2. The red line refers to the assay performed with HAO crude extract for reaction 1 and ACMSD crude extract for reaction 2. For each sample, ACMSD formation was monitored every 20s for 1500 s. Values are the average of three replicates.

5.3.2.3 Test on different substrates

According to the proposed ACMSD mechanism (Figure 5.14A, Page 80), there was no obvious reason why mesaconic or citraconic acid could not be decarboxylated. However, ACMSD was not active

on these substrates. Two different explanations may be given. The first is that neither mesaconic nor citraconic acid are able to bind into the active site. This explanation seems unlikely since several compounds presenting a dicarboxylic moiety similar to the two acids were already reported to be able to function as competitive inhibitors (Figure 5.16, Page 83) (Fukuwatari *et al.*, 2004a; Fukuwatari *et al.*, 2004b; Fukuwatari *et al.*, 2002; Huo *et al.*, 2013; Shibata *et al.*, 2007). A second explanation is that mesaconic and citraconic acid are able to bind at the catalytic cleft, but the enzyme, for some reason, is not able to catalyse their decarboxylation.

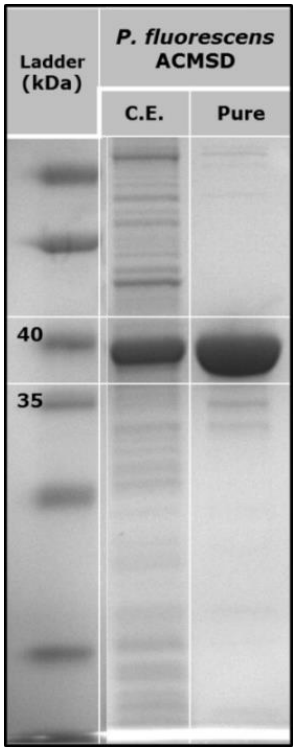


Figure 5.18: Purification of *P. fluorescens* ACMSD.
SDS-PAGE analysis of *P. fluorescens* ACMSD in crude extract (C.E.), and after purification on a Ni Sepharose column (Pure). The level of migration corresponding to ACMSD is marked between white lines.

To investigate this further *P. fluorescens* ACMSD was also tested on itaconic and fumaric acid (10 mM), two substrates that present a similar dicarboxylic acid conformation to that of citraconic and mesaconic acid. The assays were performed using both the crude extract containing the enzyme and purified ACMSD. To purify *P. fluorescens* ACMSD, the enzyme was expressed with an N-terminal

his-tag (6xHis) and then purified on a Ni Sepharose column. The protein was at least 95% pure (Figure 5.18). Substrate consumption and product formation were analysed by HPLC. No activity was observed for any of the acids in any assay format, even when zinc was added to the reactions to enhance activity (Li *et al.*, 2005; Martynowski *et al.*, 2006).

5.3.2.4 Inhibition studies

Although none of the acids tested was decarboxylated, the effect of their presence on ACMSD activity was investigated. The activity was tested upon enzymatic production of ACMS from 3-HAA, as described above. Mesaconic, citraconic, itaconic and fumaric acid were also added to the reaction mixtures and their effect on ACMS consumption rate analysed. All dicarboxylic acids inhibited ACMSD both in the crude extract and with the pure enzyme (Table 5.4). The enzyme was inhibited more when used in pure form than in crude extract, but the level of inhibition followed the same trend. The lowest inhibition was given by citraconic acid, whereas the highest inhibition was given by mesaconic acid (up to 92% inhibition). Itaconic and fumaric acid were in the middle, with fumaric acid having the highest inhibition amongst the two.

Table 5.4: Activity of ACMSD in the presence of dicarboxylic acids [a]

	ACMSD activity (%)	
	Crude extract	**Pure enzyme**
Citraconic acid	45 ± 3	20 ± 1
Itaconic acid	22 ± 1	13 ± 1
Fumaric acid	18 ± 2	10 ± 1
Mesaconic acid	14 ± 2	8 ± 2

[a] *P. fluorescens* ACMSD was assayed on *α*-amino-*β*-carboxymuconate-*ε*-semialdehyde (ACMS) in the presence of citraconic, itaconic, fumaric and mesaconic acid (4 mM). ACMS was first produced from 3-hydroxyanthranilic acid (3-HAA) using 3,4-hydroanthranilic acid dioxygenase (HAO) crude extract . ACMSD crude extract (1 mg/mL total protein) or pure ACMSD (100 µg/mL) were then added. ACMS consumption was monitored at 360 nm every 20s for 1500 s. The reaction was performed in 25 mM HEPES buffer (pH 7) with 5% glycerol at 25°C. The values of activity are in percentage compared to the activity of ACMSD in the absence of the inhibitors (0.43 $nmol \cdot s^{-1} \cdot mg_{total\ protein}^{-1}$ in crude extract and 1.97 $nmol \cdot s^{-1} \cdot mg^{-1}$ with pure ACMSD). Values are the average of three replicates and the means and errors are indicated.

In order to investigate the type of inhibition, ACMSD was tested in the presence of mesaconic acid at different concentrations. Mesaconic acid was chosen because it was the strongest inhibitor amongst the compounds tested. The acid showed competitive inhibition kinetics (Figure 5.19), with an inhibition constant (K_i) of 108±9 µM, indicating that mesaconic acid binds to the active site of ACMSD. However, the enzyme is not able to catalyse its decarboxylation. As for the other acids, inhibition studies were not performed for lack of time, but it is highly likely that inhibition occurs in a competitive manner as well.

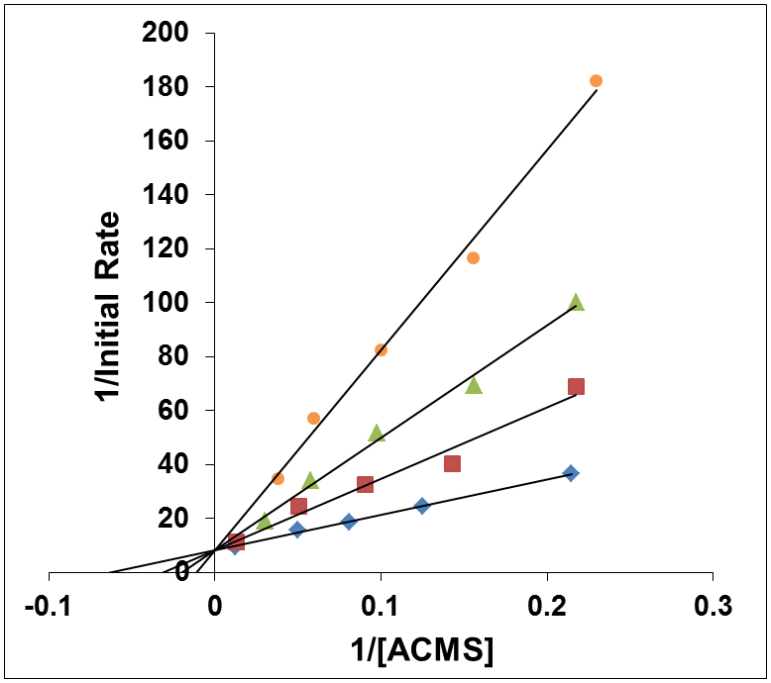


Figure 5.19: Analysis of mesaconic acid inhibition on *P. fluorescens* ACMSD. The Lineweaver-Burk plot indicates kinetics of ACMSD without mesaconic acid (light blue diamonds) and in presence of 100 µM (red squares), 200 µM (green triangles) and 500 µM (orange circles) mesaconic acid. The inhibition constant (K_i) was calculated using the competitive inhibition model of Graphpad Prism version 6.04.

Table 5.5: Kinetic parameters of *P. fluorescens* ACMSD on ACMS

K_m (µM)	Specific activity (nmol · min^{-1} · mg^{-1})	K_{cat} (s^{-1})	K_{cat}/K_m (s^{-1} · μM^{-1})
17.9±1	1536±36	1	0.06

[a] Activity values refer to the rate of ACMS consumption monitored at 360 nm. The assay was performed at 28°C in 50 mM HEPES buffer (pH 7), with the addition of glycerol (5%) and 5 µg of pure enzyme. Assays were done in triplicate and the means and errors for the kinetic constants are shown.

The kinetic data obtained for ACMSD on ACMS (Table 5.5) were slightly different from previous reports (Li *et al.*, 2006; Martynowski *et al.*, 2006). A K_m of 17.9 µM was obtained, similar to the K_m reported by Li *et a*l. (16.7 uM), whereas Martynowski *et al*. showed a lower value (9.6 µM). In regard to the specific activity, the value obtained in these experiments (1536 nmol·min^{-1}·mg^{-1}) was also lower than reported previously (4400 and 6000 nmol·min^{-1}·mg^{-1}) (Li *et al.*, 2006; Martynowski *et al.*, 2006). However, in those studies the activity of ACMSD was maximised with long incubation times in the presence of zinc, whereas in the experiments performed in this study this was not performed.

5.3.3 Discussion

P. fluorescens α-amino-β-carboxymuconate-ε-semialdehyde (ACMSD) decarboxylase was tested as a candidate for the decarboxylation of mesaconic and citraconic acid. ACMSD was expressed in *E. coli* without the need of optimisation whereas for HAO, an enzyme needed to produce the natural substrate of ACMSD, ACMS, growth and expression temperature had to be lowered to 18°C to obtain a sufficient amount of soluble protein. HAO, in higher species, is known to be unstable in solution (Koontz and Shiman, 1976) and this may explain the solubility issues observed in this study. The natural activity of ACMSD was confirmed in an assay coupled with HAO (Figure 5.17), but the enzyme was not able to decarboxylate the target substrates, citraconic and mesaconic acid, to form MAA. Citraconic and mesaconic acid present a similar dicarboxylic moiety to that of ACMS, but they lack the amino and the aldehyde moieties (Figure 5.15, Page 82). Nevertheless, the proposed mechanism of ACMSD does not indicate that the amino and aldehyde moieties are directly involved in the reaction (Figure 5.14A), and the total absence of activity on mesaconic and citraconic acid was unexpected. ACMSD was also tested on itaconic and fumaric acid, compounds characterised by a

dicarboxylic acid moiety similar to that of ACMS, but no activity was observed even with these substrates.

Further investigation showed that all the substrates tested act as inhibitors for ACMSD (Table 5.4). Mesaconic acid, the strongest inhibitor, is a competitive inhibitor (Figure 5.19), confirming that the compound binds into ACMSD's active site. Therefore, the differences between mesaconic acid and ACMS play a crucial role in the ability of the enzyme to catalyse decarboxylation. As already hypothesised, it is possible that this merely depends on thermodynamic issues, as both extra moieties could energetically favour the reaction mechanism. This problem could potentially be overcome with a protein engineering strategy aimed to modify the active site residues surrounding the catalytic centre of the substrate and, if successful, ACMSD could potentially still be used for MAA production by decarboxylation of mesaconic and citraconic acid.

However, it may also be possible that the suggested enzymatic mechanism (Huo *et al.*, 2013) (Figure 5.14A - Page 80) is wrong and that, instead, the amino group and the aldehyde have a more direct function. An alternative reaction mechanism where the tetrahedral intermediate is formed with participation of the amino moiety was already described previously (Martynowski *et al.*, 2006) (Figure 5.14B - Page 80), but it was considered energetically unfavourable because it requires a nucleophilic attack of water on a negatively charged carboxylate group. Moreover, the latter mechanism requires a carboxylate of the substrate to bind to the zinc, whereas it has been recently suggested that the two carboxylates of ACMS bind to two different Arg residues, respectively (Huo *et al.*, 2013; Huo *et al.*, 2014). Nevertheless, since instability of ACMS prevents the analysis of the substrate-bound ACMSD structure, and since all studies are instead performed with inhibitory analogues, a definitive enzymatic mechanism for ACMSD cannot be determined. Therefore, it is still possible that amino and aldehyde moieties of ACMS are somehow directly participating in the ACMSD reaction and further investigation is

needed. If this was true, neither mesaconic nor citraconic acid can be decarboxylated by ACMSD. An interesting experiment to clarify the situation would be to test the enzyme on a wider variety of substrates carrying an amino and aldehyde moiety similar to those of ACMS (for example amino mesaconate or amino citraconate), and see if/how their presence/absence affect the activity. However, the synthesis of such compounds is hard to achieve either chemically or biologically.

All the uncertainties on the enzymatic mechanism of ACMSD led me to stop studying this enzyme. More information is needed to better understand how the decarboxylation occurs, and to understand the potential of ACMSD for the production of chemicals such as MAA. ACMSD is linked to numerous diseases in humans (Guidetti and Schwarcz, 2003; Guillemin *et al.*, 2003; Mehler *et al.*, 1964; Reinhard, 2004; Schwarcz, 2004), and a better characterisation of the enzyme could also potentially lead to the design of new inhibitors. To date, most of the reported inhibitors of ACMSD present a dicarboxylic moiety in *cis* configuration (Fukuwatari *et al.*, 2004a; Fukuwatari *et al.*, 2004b; Fukuwatari *et al.*, 2002; Huo *et al.*, 2013; Shibata *et al.*, 2007). With this study, it was demonstrated that *trans* dicarboxylic acids, mesaconic and fumaric acid, can have a strong effect as well.

5.4 Mevalonate diphosphate decarboxylase

5.4.1 Introduction

Mevalonate diphosphate decarboxylase (MVD; EC 4.1.1.33) catalyses the final step of the mevalonate (MVA) pathway (Bloch, 1965), where it produces isopentenyl pyrophosphate (IPP) from (*R*)-mevalonate-5-diphosphate (MVAPP) in an irreversible reaction dependent upon ATP (Figure 5.20). The MVA pathway is present in a diverse cohort of organisms, including higher eukaryotes, archaea, a few eubacteria, fungi and plants (Kuzuyama and Seto, 2012). IPP is a key building block for a large family of metabolites such as isoprenoids, dolichols and sterols, that are crucial for a wide variety of cell vital functions (Holstein and Hohl, 2004). Many of these molecules have found applications in medicine and agriculture or as nutraceuticals, flavours and fragrances (Gershenzon and Dudareva, 2007; Kirby and Keasling, 2009).

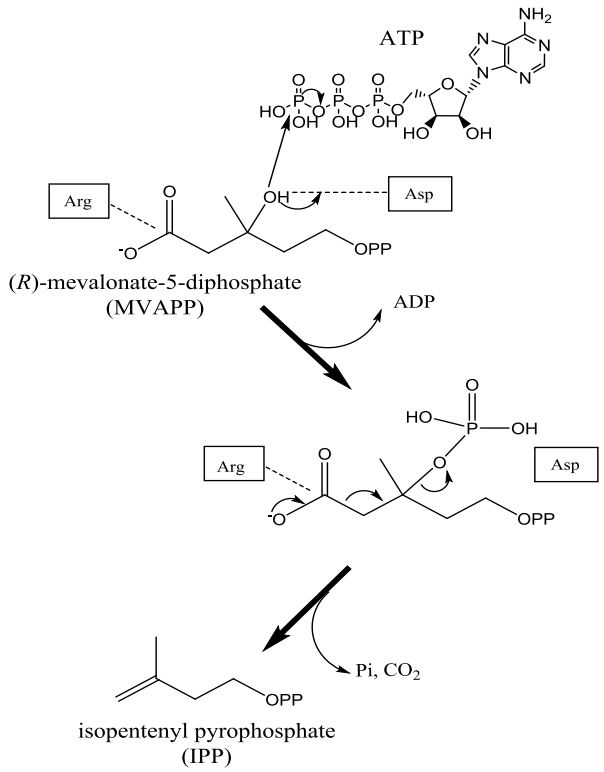


Figure 5.20: Enzymatic mechanism of MVDs

The emergence of industrial synthetic biology has led to more and more examples of metabolic engineering using enzymes from the MVA pathway, including MVD, for sustainable production of bioactive and industrial chemicals (Misawa, 2011). Products include amorpha-4,11-diene (the precursor to artemisinin; (Dahl *et al.*, 2013; Yuan and Ching, 2014)), limonene and β-carotene (Alonso-Gutierrez *et al.*, 2013; Jang *et al.*, 2011; Kim *et al.*, 2009), a range of sesquiterpenes with potential in cosmetics, pharmaceuticals and jet fuels (Scalcinati *et al.*, 2012; Wang *et al.*, 2011; Zhu *et al.*, 2014), and other biofuels such as prenol, isoprenol or other C5 compounds (Gupta *et al.*, 2014; Zheng *et al.*, 2013). Moreover, due to the importance of this route for cell viability, inhibition of the MVA pathway is also a target for the design of antibacterial or antiparasitic drugs (Rohmer *et al.*, 2004).

For all these reasons, extensive genetic, structural, inhibition and mechanistic studies from mammalian (Alvear *et al.*, 1982; Qiu *et al.*, 2007; Qiu and Li, 2006; Reardon, 1987; Voynova *et al.*, 2008), yeast (Bonanno *et al.*, 2001; Cordier *et al.*, 1999b; Krepkiy and Miziorko, 2005; Weerasinghe and Dassanayake, 2009) bacterial (Barta *et al.*, 2012; Barta *et al.*, 2011; Lefurgy *et al.*, 2010; Reuther *et al.*, 2010) and protist MVDs (Byres *et al.*, 2007) are available. MVD was classified as part of the ATP-dependent family of galactokinase, homoserine kinase, mevalonate kinase and phosphomevalonate kinase (GHMP kinase). The enzymatic mechanism (Figure 5.20) requires the presence of two catalytic residues, Asp and Arg, strictly conserved in MVDs (Barta *et al.*, 2012; Bonanno *et al.*, 2001; Byres *et al.*, 2007; Voynova *et al.*, 2008). Asp is needed to correctly position MVAPP, while functioning as a catalytic base to abstract the acidic proton found on its C3-hydroxyl. This deprotonation facilitates an in-line transfer of the γ-phosphoryl from the ATP donor, which is brought to close proximity by a rearrangement of the enzyme after binding. Phosphorylation produces an unstable intermediate that serves to enhance the leaving ability of the substituent on C3, leading to subsequent

decarboxylation to proceed. After decarboxylation, structural changes can reverse to allow release of the products ADP and IPP.

The binding of the substrates, ATP and MVAPP, requires the presence of specific residues, which are also highly conserved within MVDs. The residues involved in ATP interactions are situated mostly in a domain that is present in all GHMP kinase family enzymes (Bonanno *et al.*, 2001) but there is also involvement of other amino acids (Byres *et al.*, 2007). The residues suggested to interact with MVAPP are not reassembled in one specific region but are instead spread along the whole length of the amino acid sequence. Some of these residues are specific for the C1 carboxylate of MVAPP, helping in positioning the substrate and subsequently driving its decarboxylation (Barta *et al.*, 2012; Byres *et al.*, 2007; Voynova *et al.*, 2008). Other residues are instead suggested to interact more specifically with the diphosphate moiety of MVAPP (Barta *et al.*, 2011), and more amino acids contribute to generate a hydrophobic region that allows proper folding of the active site around MVAPP (Byres *et al.*, 2007).

What brought this enzyme to the attention is the ability to couple dehydration and decarboxylation of MVAPP *via* initial phosphorylation of the C3-hydroxyl (Barta *et al.*, 2012; Byres *et al.*, 2007; Voynova *et al.*, 2008). This energetically favourable solution allows the dehydration of alcohols to produce olefins, which is otherwise unfavourable. Among the acids identified as potentially convertible to MAA, citramalic acid shows the same overall characteristics and chemical structure as MVAPP (Figure 5.21), with a hydroxyl substituent at C3 and a carboxylate at C1. In order for citramalic acid to be converted to MAA, a dehydration/decarboxylation must occur, exactly as the one catalysed by MVD.

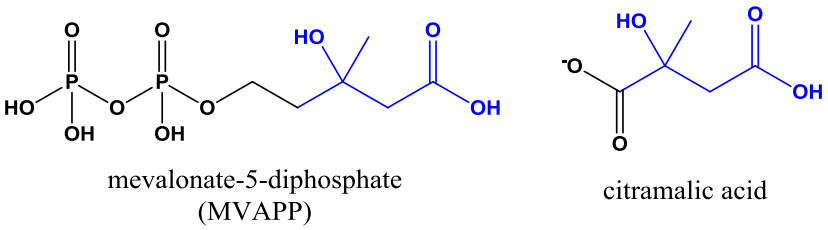


Figure 5.21: Structure similarity of citramalic acid with MVAPP

Although citramalic acid presents similarities to MVAPP, there is also a significant difference between the two chemicals. MVAPP carries a long tail with a terminal diphosphate moiety, whereas citramalic presents a shorter carboxylate substituent (Figure 5.21). Since MVD has evolved to bind the diphosphate moiety of MVAPP, the difference with citramalic acid could be crucial. The presence of a negatively charged, potentially reactive, polar carboxylate could spoil both substrate specificity/positioning and catalytic efficiency of MVD. Inhibition studies on rat liver MVD using fluorine substituents at carbons close to the one bearing the hydroxyl group (C3 and C4) showed that the rate of phosphorylation can be reduced (Reardon, 1987). However, since there is no direct evidence that the conversion of citramalic acid to MAA cannot work, these considerations did not decrease the intention to test the enzyme but rather increased the awareness of a possible protein engineering strategy to adapt MVD to accept citramalic acid's carboxylate.

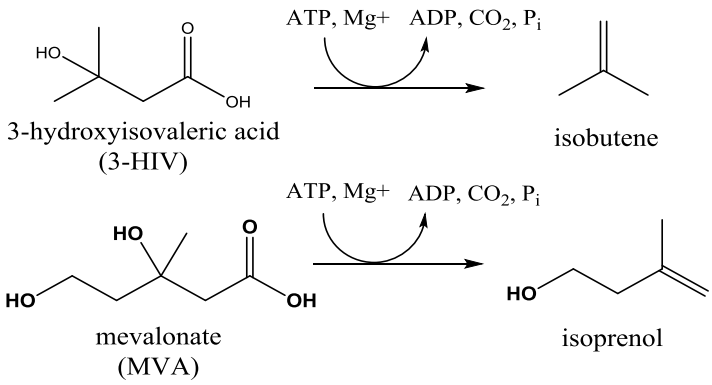


Figure 5.22: Substrate promiscuity of MVDs

The interest in the enzyme was further strengthened by the ability of MVD to act on other substrates similar to MVAPP but without the diphosphate moiety. In particular, MVD was able to convert 3-hydroxyisovaleric acid (3-HIV) to isobutene and MVA to isoprenol (Figure 5.22) (Gogerty and Bobik, 2010; Marliere *et al.*, 2011). These examples demonstrated promiscuity of MVDs towards non-natural substrates.

Among different MVDs, the homologue from *Picrophilus torridus* was reported to catalyse formation of isobutene and isoprenol more efficiently than others (Marliere, 2010; Marliere *et al.*, 2011), suggesting a higher level of substrate promiscuity. Therefore, this enzyme was the first choice to test MAA formation from decarboxylation/dehydration of citramalic acid. However, since *P. torridus* MVD was still poorly characterized, Iit wasalso decided to test *S. cerevisiae* MVD, which has been more extensively studied (Bonanno *et al.*, 2001; Cordier *et al.*, 1999a; Gogerty and Bobik, 2010; Krepkiy and Miziorko, 2004; Krepkiy and Miziorko, 2005; Weerasinghe and Dassanayake, 2009).

5.4.2 Results

5.4.2.1 Enzyme expression

Genes encoding both enzymes were synthesized and optimized for expression in *E. coli* (Appendix 9.1, Page 204-205). The optimised gene sequence encoding for *P. torridus* MVD was taken from a study where soluble expression in *E. coli* BL21 was significantly improved by such optimisation (Marliere, 2010). After insertion of the genes into the expression vector and after transformation of *E. coli*, overexpression was analysed. Information regarding successful expression of *S. cerevisiae* MVD in *E. coli* (Bonanno *et al.*, 2001; Gogerty and Bobik, 2010; Krepkiy and Miziorko, 2005) was used, even though the *E. coli* strain and the plasmid were different. Optimal expression was obtained at 37°C induction for 5-24 h (Appendix 9.3 – Page 216). Expression of *P. torridus* MVD was

tested at different temperatures. Although consistently good, its level of production reached optimal levels at 37°C induction for 5-24 h (Appendix 9.3 – Page 215), like the *S. cerevisiae* counterpart. The estimated molecular weight of *P. torridus* MVD was around 37 kDa and that of *S. cerevisiae* MVD was 44 kDa, consistent to previous reports (Gogerty and Bobik, 2010; Marliere, 2010). Both enzymes were expressed at larger scale to obtain crude extracts for the enzymatic assays.

5.4.2.2 Enzyme activity

Crude extracts were mixed with 3-HIV and formation of isobutene monitored by GC-MS. *S. cerevisiae* MVD was tested at pH 7.5 and 30°C, and *P. torridus* MVD at pH 5.5 and 50°C, conditions reported to be optimal, respectively (Gogerty and Bobik, 2010; Marliere, 2010). Background formation of isobutene was observed in all the controls (Table 5.6), probably due to the known spontaneous degradation of 3-HIV (Marliere, 2010; Pressman and Lucas, 1940). However, only samples containing the enzymes of interest and the cofactor, ATP, showed an increased level of isobutene production (Table 5.6). This confirms that both catalysts were expressed in active form. Formation of isobutene was significantly higher with *P. torridus* MVD (Table 5.6), as already reported previously (Marliere, 2010).

Table 5.6: *S. cerevisiae* and *P. torridus* MVDs enzyme activity confirmation [a]

Enzyme	Only buffer ($pmol \cdot min^{-1}$)	Empty	MVD No ATP	MVD
		($pmol \cdot min^{-1} \cdot mg_{(total\ protein)}^{-1}$)		
P. torridus MVD	0.5 ± 0.01	0.5 ± 0.02	0.6 ± 0.01	261 ± 19
S. cerevisiae MVD	0.5 ± 0.01	0.6 ± 0.1	0.6 ± 0.3	1.6 ± 0.3

[a] *S. cerevisiae* and *P. torridus* MVD were assayed on 3-HIV (50 mM) as substrate. ATP (40 mM) was also added when not specifically indicated. The reaction was performed in 50 mM Tris buffer (pH 7.5) at 30°C for *S. cerevisiae* MVD and in 50 mM acetate buffer (pH 5.5) at 50°C for *P. torridus* MVD. To both buffers, KCl (20 mM) and $MgCl_2$ (10 mM) was added. The reaction was performed for 96 h with gentle shaking. Results shown represent the rates of isobutene formation over 48h, since production was not linear after that. (Only buffer) refers to the assay performed in the absence of crude extract. (Empty) refers to the assay performed with crude extract from *E. coli* containing empty vector. (MVD - No ATP) refers to the assay performed with crude extract containing the enzymes of interest but no ATP. (MVD) refers to the assay performed with crude extract containing the enzymes of interest. Values of isobutene formation are the average of two replicates.

P. torridus and *S. cerevisiae* MVDs were then tested on both citramalic acid enantiomers, (*R*) and (*S*). In principle, (*S*)-citramalic acid should be the favoured substrate, since it has the same configuration as MVAPP. However, at this stage, it was preferred to test both forms. Consumption of citramalic acid and formation of MAA were monitored by HPLC. Unfortunately, neither citramalic acid consumption nor MAA formation was observed. Comparison with chromatograms from controls did not show any anomalous peak that could have represented the formation of intermediate products.

5.4.2.3 Does citramalic acid bind to MVDs?

Table 5.7: Isobutene formation rates in the presence of citramalic acid [a]

	***P. torridus* MVD** ($pmol \cdot min^{-1} \cdot mg_{(total\ protein)^{-1}}$)	***S. cerevisiae* MVD** ($pmol \cdot min^{-1} \cdot mg_{(total\ protein)^{-1}}$)
No citramalic acid	261 ± 19	1.6 ± 0.3
(*R*)-citramalic acid	255 ± 7	1.6 ± 0.1
(*S*)-citramalic acid	253 ± 19	1.7 ± 0.7

[a] Both 3-HIV (50 mM) and ATP (40 mM) were used in the assay. The reaction was performed in 50 mM Tris buffer (pH 7.5) at 30°C for *S. cerevisiae* MVD and in 50 mM acetate buffer (pH 5.5) at 50°C for *P. torridus* MVD. To both buffers, KCl (20 mM) and $MgCl_2$ (10 mM) was added. The reaction was performed for 96 h with gentle shaking. Results shown represent the rates of isobutene formation calculated over 48h, since production was not linear after that. (*R*)- or (*S*)-citramalic were added at concentration of 5 mM. Assays were done in duplicate and the means and errors for formation rates are shown.

In order to understand whether or not citramalic acid binds to the active site of MVDs, its potential inhibition on the target enzymes was analysed. Both (*S*) and (*R*)-citramalic acid were tested. Crude extract containing *P. torridus* and *S. cerevisiae* MVD was prepared from engineered *E. coli* cells. The extract was then mixed with 3-HIV in the absence and presence of the two citramalic acid enantiomers. The experiments were performed in sealed vials so that isobutene could be measured by GC-MS by analysis of the headspace. Since isobutene can form spontaneously from 3-HIV (Pressman and Lucas, 1940), controls with no enzyme were always compared and the rate of isobutene formation in controls was subtracted from that of the enzymatic reactions. No significant differences were observed for the formation rates of isobutene in the presence and absence of the acids (Table 5.7). Therefore, *P. torridus* and *S. cerevisiae* MVD are

not inhibited by citramalic acid. This suggests that citramalic acid may not bind to the active site and that this could potentially be the reason why conversion to MAA does not occur.

5.4.3 Discussion

S. cerevisiae and *P. torridus* mevalonate diphosphate decarboxylases (MVD) were tested for the decarboxylation of citramalic acid. Expression of both enzymes in *E. coli* did not require particular optimisation. Confirmation of their expression as active catalyst was performed by monitoring the formation of isobutene from 3-HIV (Gogerty and Bobik, 2010; Marliere *et al.*, 2011) (Table 5.6). However, the enzymes were not able to convert either (*S*)- or (*R*)-citramalic acid to MAA.

Expectations for the use of *S. cerevisiae* MVD and *P. torridus* MVD were high. First of all, there is a significant similarity between citramalic acid and MVAPP, the natural substrate (Figure 5.21). Moreover, both enzymes had shown the ability to catalyse decarboxylation/dehydration of 3-HIV, a substrate with a structure similar to MVAPP but lacking the diphosphate moiety (Gogerty and Bobik, 2010; Marliere, 2010; Marliere *et al.*, 2011). Nevertheless, citramalic acid carries a polar carboxylate moiety at C3 that is not present in any of the known substrates. MVDs can act on substrates with non-polar substitutions at C3 (Lefurgy *et al.*, 2010), but adverse effect with substrate with polar substitutions have already been reported (Reardon, 1987). Further investigations suggest that citramalic acid does not even bind to the active site of MVD, since isobutene formation was not inhibited in the presence of the substrate (Table 5.7). Therefore, the presence of a negatively charged, polar carboxylate may prevent binding of citramalic acid and, subsequently, impede the reaction.

All the selected enzyme candidates did not catalyse formation of MAA from either itaconic, mesaconic, citraconic or citramalic acid. Since these enzymes were already considered to be among the best

ones for the purpose, it was decided that, rather than testing other decarboxylases, a protein engineering attempt could be a more promising approach to develop the ability to catalyse the target reactions, since protein engineering can be used to develop unprecedented enzyme catalysis (Bornscheuer and Kazlauskas, 2004). However, time did not permit engineering all the enzymes initially selected, and it was decided to focus on *S. cerevisiae* and *P. torridus* MVD only. MVDs catalyse the dehydration/decarboxylation of hydroxyalkanoic acids, as required for the conversion of citramalic acid to MAA (Figure 5.23), and have shown substrate promiscuity (Gogerty and Bobik, 2010; Marliere *et al.*, 2011), good starting point to develop new substrate specificities (Bornscheuer and Kazlauskas, 2004). Moreover, the interest in this enzyme, and in the conversion of citramalic acid to MAA, was strengthened by the recent development, in our group, of a process for the production of citramalic acid in high concentrations using engineered *E. coli* (Eastham *et al.*, 2015). Coupling this process with the conversion to MAA using MVD would be the quickest way of having a process for the bioproduction of MAA.

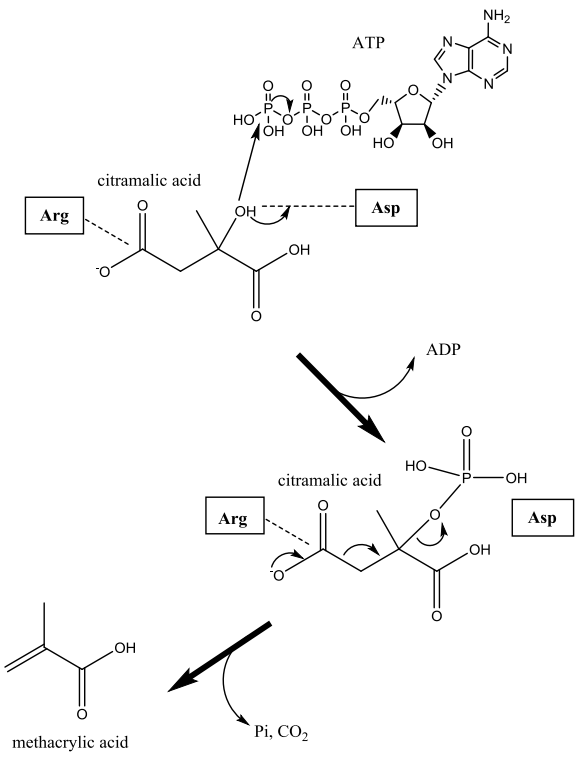


Figure 5.23: Potential conversion of citramalic acid to MAA catalysed by MVDs

5.5 Protein engineering on MVDs

5.5.1 Introduction

In the previous chapter, it was shown that citramalic acid does not bind to the active site of MVDs. Thus, in regards to a protein engineering strategy aimed to enable MVDs to catalyse the decarboxylation of this substrate, this problem must be overcome, and the active site, and in particular the binding pocket, must be modified. In this chapter, the procedures and techniques that allowed to perform protein engineering on these enzymes are described. At first, a screening assay to detect enzyme activity in microtiter plate format was developed and optimised. The throughput level of the screening protocol was low and, in order to reduce screening efforts without compromising the likelihood of evolving enhanced enzymes, variation was introduced following a Iterative Saturation Mutagenesis (ISM) approach (Reetz and Carballeira, 2007). Selection of residues was possible thanks to the great level of information available on MVDs, which includes mechanistic, genetic and structural studies (Alvear *et al.*, 1982; Barta *et al.*, 2012; Barta *et al.*, 2011; Bonanno *et al.*, 2001; Byres *et al.*, 2007; Cordier *et al.*, 1999b; Krepkiy and Miziorko, 2005; Lefurgy *et al.*, 2010; Qiu *et al.*, 2007; Qiu and Li, 2006; Reardon, 1987; Reuther *et al.*, 2010; Voynova *et al.*, 2008; Weerasinghe and Dassanayake, 2009).

Libraries were then generated using the QuickChange Lightening Site-Directed Mutagenesis Kit, which is based on the principle of whole plasmid amplification with mutagenic primers (Papworth *et al.*, 1996). This kit allowed for quick and efficient simultaneous saturation mutagenesis at the decided positions. Libraries were then transformed into *E. coli* BL21(DE3)pLysS and clones screened for the target activity.

5.5.2 Results

5.5.2.1 Development and optimization of a screening assay

The target reaction to screen for was the ATP-dependent decarboxylation/dehydration of citramalic acid (Figure 5.23, Page 100). This would consist of a first phosphorylation at the hydroxyl on C3 of the substrate, subsequent dephosphorylation and concomitant decarboxylation (Barta *et al.*, 2012; Bonanno *et al.*, 2001; Byres *et al.*, 2007; Voynova *et al.*, 2008). This reaction would consume citramalic acid and ATP and would release MAA, ADP, inorganic phosphate and CO_2 (Figure 5.23, Page 100). The monitoring of any of these parameters could potentially be exploited to develop a screening assay. In the next sections, the evaluation of different techniques to monitor these parameters in 96-well plate format is described. The goal was to develop a robust, reliable and reproducible screening method to be used in protein engineering procedures. *In vitro* screening procedures which utilised crude extracts containing the enzymes were considered more appropriate than *in vivo* selection procedures using living cells, because of problems with citramalic acid uptake into cells, MAA release from cells for detection, MAA toxicity (Eastham *et al.*, 2015), and to the impossibility to develop a complementation system which could be linked to the survival ability of the cells.

5.5.2.1.1 Enzyme expression and cell lysis in 96-well plate format

A good procedure for enzyme expression and cell lysis must be developed in 96-well plate format. Therefore, expression of *P. torridus* MVD was tested in 96-well plates using engineered *E. coli* cells. Cells not expressing the enzyme of interest were also tested as a control (referred in the text as empty vector crude extract samples). Single colonies were inoculated into individual wells and enzyme expression was induced overnight under the same conditions developed during expression optimization (LB medium,

37°C). Cell pellets were then collected by centrifugation and lysed to obtain the crude extracts. Three different lysis protocols were analysed. The difference between the three was the composition of the solution used for lysis: one contained lysozyme from chicken egg white (LysChick), one contained rLysozyme™ Solution (rLys) and one contained BugBuster® Reagent without any lysozyme (BugBus). Enzyme expression was evaluated by SDS-PAGE and the efficiency of the lysis method by calculating the total protein concentration.

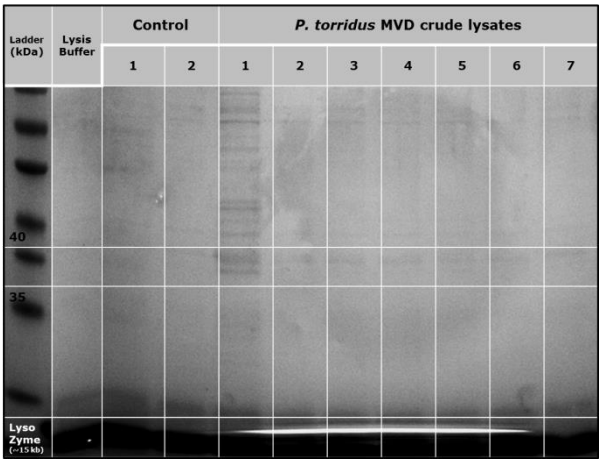
Table 5.8: Efficiency of different cell lysis procedures in 96-well plate format [a]

Lysis solution	**Control** [total protein] (mg/mL)	***P. torridus* MVD crude extracts** [total protein] (mg/mL)
LysChick	0.349 ± 0.118	0.508 ± 0.071
rLys	0.201 ± 0.052	0.281 ± 0.039
BugBus	0.640 ± 0.047	0.700 ± 0.110

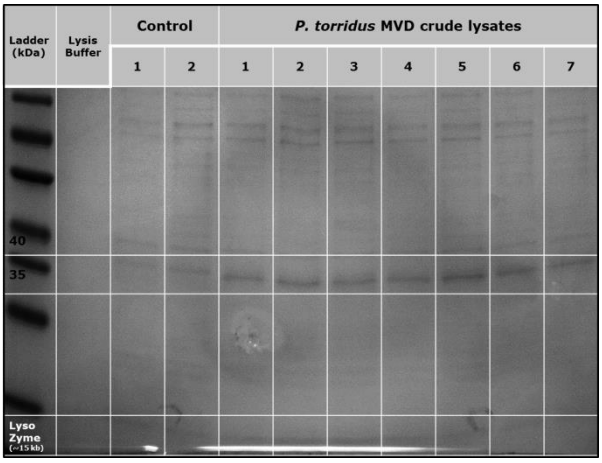
[a] Values are the average of randomly tested samples. Control refers to empty vector crude extract.

In all three cases, the expression level of *P. torridus* MVD was good (Figure 5.24). When using LysChick, the concentration of lysozyme was very high compared to the rest of the proteins (Figure 5.24A). This would interfere with measuring the concentration of the target enzymes. The situation improved when using rLys (Figure 5.24B), since less lysozyme was needed. However, among the three methodologies tested, the one employing BugBus gave the highest total net concentration of protein for both the control (0.640 mg/mL) and the cells expressing *P. torridus* MVD (0.700 mg/mL) (Table 5.8). Therefore, this lysis technique was used in all following experiments. The same expression and lysis procedures were then tested for *S. cerevisiae* MVD. The results were as good as for *P. torridus* MVD (Figure 5.25), with an average total protein concentration of 0.737. This indicated that the protocol could be used for both enzymes.

A



B



C

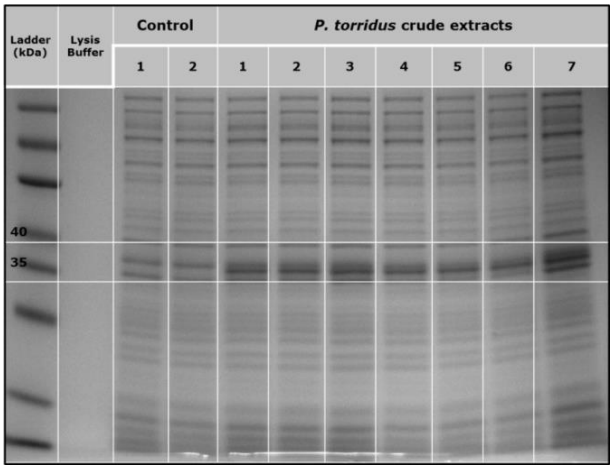


Figure 5.24: SDS-PAGE analysis of crude extracts prepared in 96-well plate format
Expression level of *P. torridus* MVD was analysed. Three different lysis solutions (LysChick (A), rLys (B) and BugBus (C)) were used and compared. (Lysis buffer) refers to a sample containing only the respective lysis solution. (Control) refers to two random samples containing empty vector crude extract. (*P. torridus* MVD crude extracts) refers to seven random samples containing extracts obtained from cells expressing *P. torridus* MVD. The level of migration corresponding to *P. torridus* MVD is marked between white lines. In (A) and (B), the level of migration corresponding to the lysozyme is labelled as "Lysozyme". All the other bands correspond to other proteins present in the extracts.

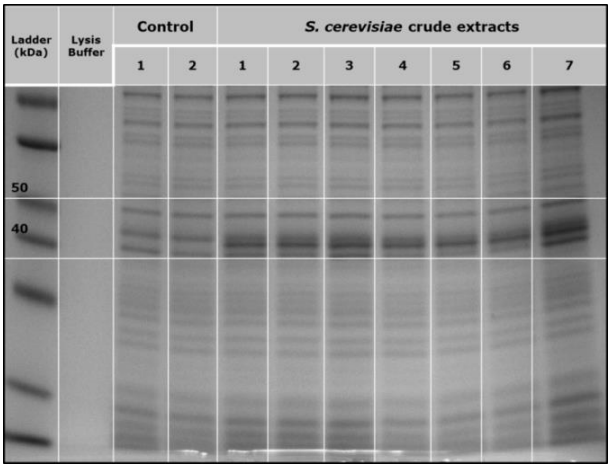


Figure 5.25: Analysis of expression of *S. cerevisiae* MVD in 96-well plate
S. cerevisiae MVD crude extracts were obtained using BugBus solution. (Lysis buffer) refers to a sample containing the lysis solution only. (Control) refers to two random samples containing empty vector crude extract. (*S. cerevisiae* MVD crude extracts) refers to seven random samples containing extracts obtained from cells expressing *S. cerevisiae* MVD. The level of migration corresponding to *S. cerevisiae* MVD is marked between white lines. All the other bands correspond to other proteins present in the extracts.

5.5.2.1.2 MAA formation/citramalic consumption

MAA formation and citramalic acid consumption are key parameters that could be monitored to detect activity of MVDs on citramalic acid. MAA detection could potentially be done by UV analysis, as MAA has unique optical properties with a characteristic peak between 205 and 210 nm (Fikhtengol'ts *et al.*, 1969). Therefore, the MAA spectrum was compared with that of citramalic acid and ATP (Figure 5.26), substrate and cofactor of the reaction, respectively. ADP was not tested because it was not considered crucial as it is produced only after hydrolysis of ATP. Unfortunately, the ATP spectrum showed a strong absorbance at around 210 nm, overlapping with MAA (Figure 5.26). Therefore, this methodology was abandoned.

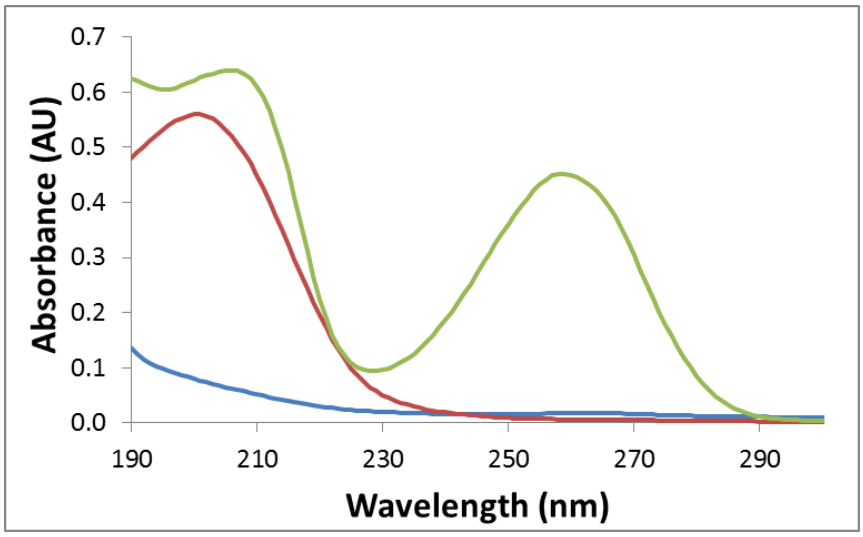
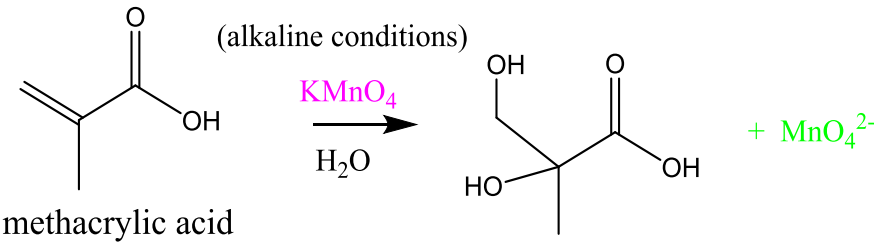


Figure 5.26: Absorbance spectrum of MAA, citramalic acid and ATP
Citramalic acid, MAA and ATP were prepared in mQ water at a concentration of 5 mM. A 100-fold dilution was then prepared in quartz cuvettes and absorbance spectrum obtained using an Agilent 8453 UV-visible spectrophotometer. The blue line refers to citramalic acid, the red line to MAA and the green line to ATP.

Another way to monitor MAA formation would be to react it with potassium permanganate ($KMnO_4$) or bromine (Br_2). $KMnO_4$ and Br_2 can react with the double bond present on MAA by hydroxylation or halogenation of the alkenyl moiety, respectively (Figure 5.27). Both reactions are characterised by a colour change. Alkenes react with $KMnO_4$ to give the respective diol (Wiberg and Saegebarth, 1957). $KMnO_4$ in solution is purple and, in neutral or slightly alkaline conditions, it first becomes green/yellow and then produces a dark brown precipitate. This is due to the fact that permanganate ions (MnO_4^-) are first reduced to green manganate ions (MnO_4^{2-}) and then to solid manganese dioxide (MnO_2) (Wiberg and Saegebarth, 1957). The reddish-brown Br_2 dissolved in water (bromine water) at room temperature and under acidic conditions undergoes electrophilic addition with double bonds and decolourises during the reaction depending on the concentration of reactive alkenes. Both reactions should be specific for MAA because citramalic acid does not present any alkenyl moiety.

A



B

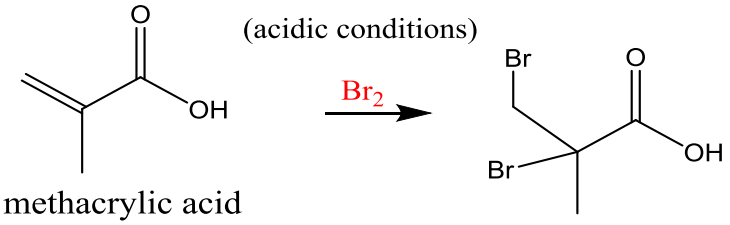
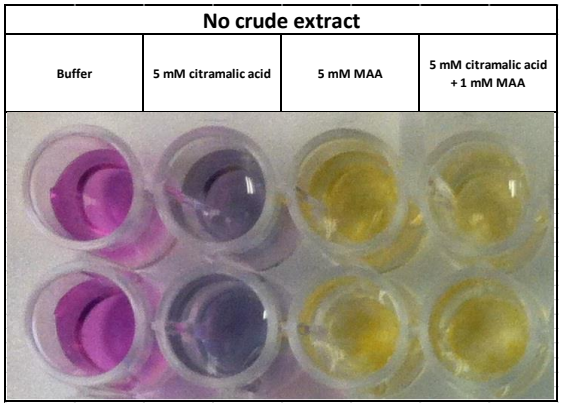


Figure 5.27: Potential hydroxylation and halogenation of MAA
Hydroxylation of MAA with potassium permanganate ($KMnO_4$) (A) and halogenation of MAA with Bromine (Br_2) (B).

A preliminary evaluation of $KMnO_4$ reactivity with different concentrations of MAA was initially performed in the presence and absence of crude extract (Figure 5.28). In the absence of crude extract, $KMnO_4$ reacted specifically with MAA when present as the sole chemical in the reaction and also when mixed with citramalic acid. No reactivity was observed with citramalic acid only and in the controls (Figure 5.28A). However, in the presence of crude extract, no clear difference was visible anymore (Figure 5.28B), as confirmed by spectrophotometric analysis at 340 nm. Therefore, $KMnO_4$ also reacted with extra components present in the crude extract, probably other alkenes or other chemicals that can be oxidised (Amin *et al.*, 2010; Chitra *et al.*, 2004; Rajendraprasad and Basavaiah, 2009).

A



B

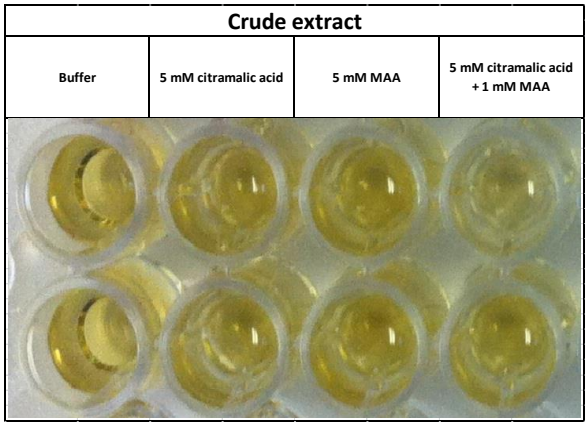
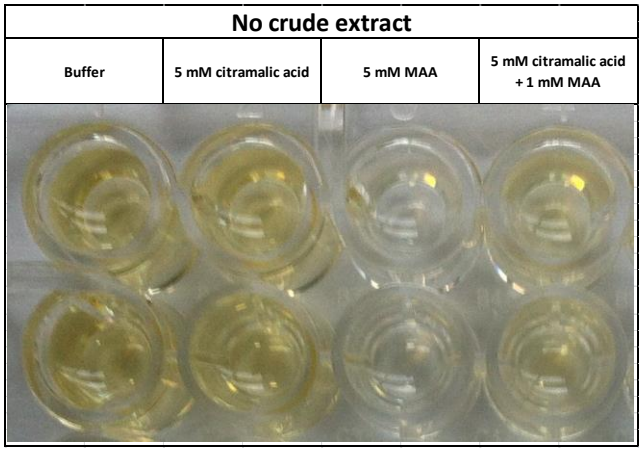


Figure 5.28: Potassium permanganate reactivity on MAA
Reaction was performed in a 96-well plate. Samples (250 µL) consisted of (B) crude extract (0.5 mg/mL total protein) prepared in 50 mM acetate buffer (pH 5.5) with addition of 10 mM $MgCl_2$ and 20 mM KCl , or (A) only buffer. 100 mM NaOH (20 µL) and 5 mM $KMnO_4$ solution (50 µL) were then added, and the reaction incubated at room temperature for 10 min. Oxidation of products was observable by decolouration of $KMnO_4$ and was monitored at 340 nm using a FLUOstar Optima plate reader (BMG Labtech). (Buffer) indicates control samples containing only buffer. The other samples contained 5 mM MAA, 5 mM citramalic acid and a mixture of 5 mM citramalic and 1 mM MAA.

As for $KMnO_4$, reactivity of Br_2 was tested in the presence and absence of crude extract (Figure 5.29). In this case, an evident decolouration depending on MAA presence/concentration was observed even in the presence of crude extract, suggesting that Br_2 is less affected by the presence of extra components. Since this looked very promising, a simulation of citramalic acid conversion to MAA was set up in a 96-well plate and analysed. This was done by mixing the two compounds at proportioned concentrations (5 mM citramalic acid or 4 mM citramalic acid and 1 mM MAA, 1 mM citramalic acid and 4 mM MAA or 5 mM MAA). ATP and crude extract

were also present. Bromine water was then added and the decolouration evaluated at 390 nm (Table 5.9).

A



B

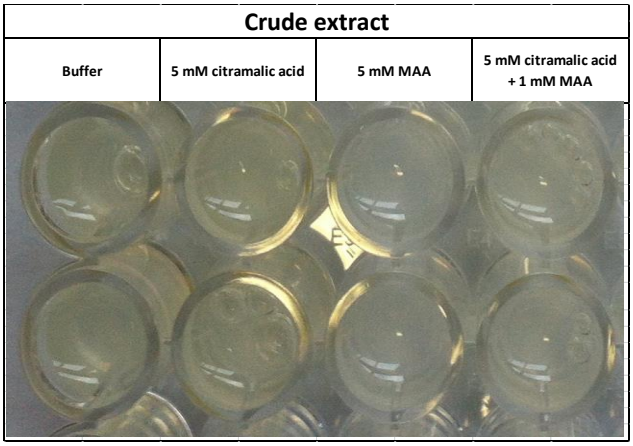


Figure 5.29: Bromine water reactivity on MAA
Reaction was performed in a 96-well plate. Samples (250 µL) consisted of (B) crude extract (0.5 mg/mL total protein) prepared in 50 mM acetate buffer (pH 5.5) with addition of 10 mM $MgCl_2$ and 20 mM KCl , or (A) only buffer. 50 mM Br_2 solution (50 µL), previously prepared by acidification of 50 mM bromide-bromate with concentrated H_2SO_4, was added and the reaction incubated at room temperature for 30 s. Halogenation of products was observable by decolouring of Br_2 solution and was monitored at 390 nm using a FLUOstar Optima plate reader (BMG Labtech). (Buffer) indicates control samples containing only buffer. The other samples contained 5 mM MAA, 5 mM citramalic acid and a mixture of 5 mM citramalic and 1 mM MAA.

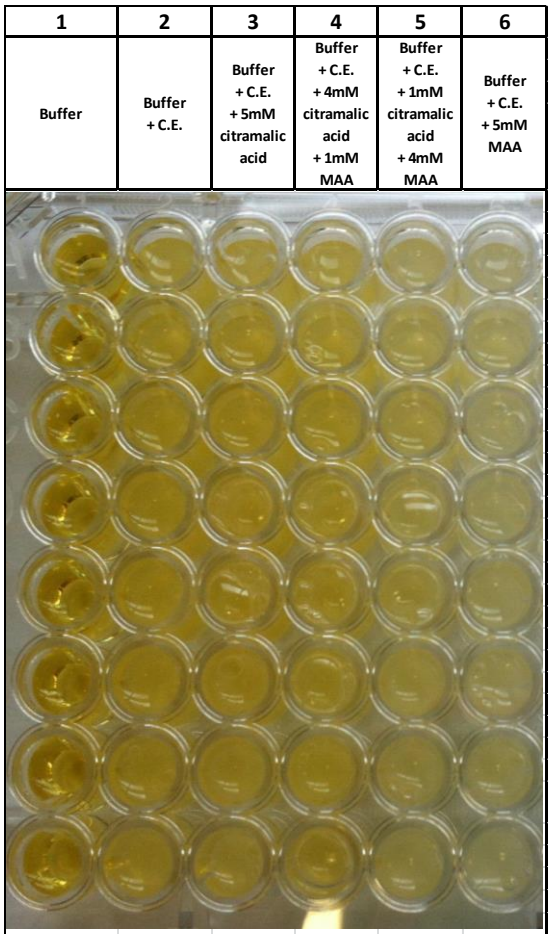


Figure 5.30: Bromine water screening assay simulation
Reaction was performed in a 96-well plate as described in Figure 5.29 but 5 mM ATP was also added to the samples. Two negative controls consisting of only buffer and buffer + crude extract (C.E.) were compared with 5 mM citramalic acid, 4 mM citramalic acid + 1 mM MAA, 1 mM citramalic acid + 4 mM MAA, 5 mM MAA.

Table 5.9: Bromine screening assay simulation (Absorbance values) [a]

Buffer	Buffer + C.E.	Buffer + C.E. + 5 mM citramalic acid	Buffer + C.E + 4 mM citramalic acid + 1 mM MAA	Buffer + C.E. + 1 mM citramalic acid + 4 mM MAA	Buffer + C.E. + 5 mM MAA
0.996±0.032	2.391±0.081	2.463±0.096	2.260±0.061	1.954±0.068	1.875±0.124

[a] Average values of absorbance at 390 nm are indicated with their respective errors.

As expected, the decolouration of Br_2 was higher with higher concentrations of MAA (Figure 5.30). The presence of citramalic acid did not have a significant effect. The absorbance value at 390 nm were consistently lower with the higher concentration of MAA (Table 5.9), making it possible to detect its formation. Although this was

very promising, some limitations were encountered. First of all, Br_2 is highly toxic. The extra precautions that must be taken when handling the chemical were of serious concern and very laborious. Moreover, the differences in Br_2 decolouration were only visible for less than 30 s, at the end of which Br_2 reacted with other crude extract components and decolourised completely in all samples. In such a short time, spectrophotometric analysis and visual observation had to be done very quickly, with the risk of spoiling results and repeatability. Attempts in improving reactivity using different concentrations of Br_2 were unsuccessful. Therefore, this methodology was considered impractical.

5.5.2.1.3 Inorganic phosphate release

MVD catalyses the conversion of ATP to ADP, and a molecule of inorganic phosphate (Pi) is released (Figure 5.23, page 100 - Figure 5.20, page 92). A common reaction used for the detection of Pi is its interaction with ammonium molybdate ($(NH_4)_6Mo_7O_{24}$) to form a phosphomolybdate complex. This complex, in the presence of a reducing agent (*e.g.* ascorbic acid), forms another complex that is coloured, called molybdenum blue (Figure 5.31). Molybdenum blue can be quantified by monitoring absorbance at 660 nm (Gawronski and Benson, 2004). The utilization of $(NH_4)_6Mo_7O_{24}$ for Pi detection in 96-well plate format has already been reported (Gawronski and Benson, 2004). More recently, a similar procedure was employed to analyse MVD activity (Marliere *et al.*, 2011), but the assays were performed with purified enzymes rather than with crude extracts.

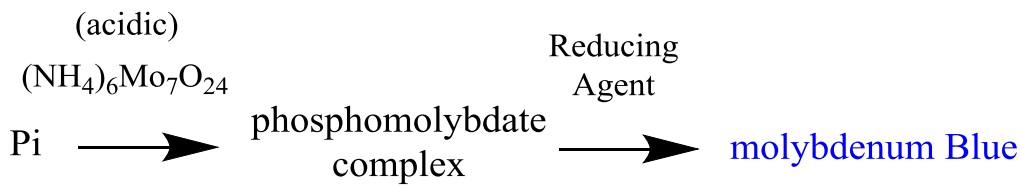


Figure 5.31: Reaction of inorganic phosphate with ammonium molybdate
Molybdenum blue is formed by interaction of inorganic phosphate (Pi) and ammonium molybdate ($(NH_4)_6Mo_7O_{24}$).

The feasibility of this approach was tested in crude extract containing *P. torridus* MVD, with 3-HIV and ATP. Formation of isobutene was monitored by GC-MS after 48 h and, upon confirmation that the reaction occurred, samples were treated with $(NH_4)_6Mo_7O_{24,}$ as described previously (Gawronski and Benson, 2004; Marliere *et al.*, 2011). *P. torridus* MVD, rather than *S. cerevisiae* MVD, was tested due to its higher activity towards 3-HIV. Controls not containing crude extract, or containing empty vector crude extracts, were also analysed. Quantification of molybdenum blue was performed by reading absorbance at 660nm. Contrary to expectations, the concentration of Pi in samples with active MVD was not higher than the controls with MVD-free crude extract (Table 5.10). This may have been due to the presence of intracellular Pi and to background activities of natural *E. coli* phosphatases. Therefore, even this approach could not be used as a screening assay in crude extracts.

Table 5.10: Inorganic phosphate concentration in ammonium molybdate test evaluation [a]

	No crude extract	**Empty vector crude extract**	***P. torridus* MVD crude extract**
Inorganic phosphate (mM)	4.13 ± 0.08	8.78 ± 0.04	7.96 ± 0.12

[a] The reaction (1 mL) was assembled in GC vials which had PTFE screw caps and contained crude extract (0.5 mg/mL total protein),prepared in 50 mM acetate buffer (pH 5.5) with addition of 10 mM $MgCl_2$ and 20 mM KCl, 5 mM 3-HIV and 5 mM ATP. As controls, samples containing only buffer and MVD-free crude extract were also analysed. The reaction was incubated at 37°C for 48 h with gentle shaking. The mixture (10 µL) was then mixed with a 0.45 M L-ascorbic acid/2.5 mM $(NH_4)_6Mo_7O_{24}$ solution (150 µL) in 1 M HCl and held at room temperature for 5 min. Reaction was terminated by adding 1 M citric acid (10 µL) and allowed to equilibrate for 15 min at room temperature. Quantification of Pi was performed by reading the absorbance at 600 nm with a FLUOstar Optima plate reader (BMG Labtech). A linear calibration curve for Pi was firstly obtained using KH_2PO_4 in concentration range of 0.25 – 10 mM. Values are the average of three replicates and the means and errors are indicated.

Already demonstrated when testing the ATPase activity of Hsp90 chaperone (Avila *et al.*, 2006), another option that could potentially be employed to analyse the release of Pi was initially considered and then abandoned because it was believed to be too laborious and expensive. This method is based on a coupled enzymatic reaction that requires the activity of three additional enzymes: maltose phosphorylase, glucose oxidase and horseradish peroxidase.

Although sensitivity can be high (Avila *et al.*, 2006), the assay is very expensive.

5.5.2.1.4 ADP release

The biological conversion of ATP to ADP could also potentially be used to monitor activity of MVD (Figure 5.23, page 100). A colorimetric method commonly used for this purpose is to couple the main enzymatic reaction with two reactions catalysed by pyruvate kinase (PK) and lactic acid dehydrogenase (LDH). Phosphoenolpyruvate (PEP) and NADH are added and then PK uses ADP as cofactor to catalyse the conversion of PEP to pyruvate, which is then converted to lactate by LDH (Figure 5.32). The latter reaction is NADH dependent and NADH is oxidised. This oxidation can be monitored spectrophotometrically at 340 nm. An extra advantage of this system is that the reaction catalysed by PK regenerates ATP that can be reutilised by MVD. This system is widely employed for studies on ATP-dependent enzymes and has been used for MVD as well (Alvear *et al.*, 1982; Barta *et al.*, 2012; Barta *et al.*, 2011; Krepkiy and Miziorko, 2005; Krepkiy and Miziorko, 2004). Although it is usually performed using purified enzymes, there are examples proving its suitability in crude extract-based assays as well (Baumann *et al.*, 2001; Brown *et al.*, 2012; Goddard and Reymond, 2004).

The screening assay was tested in 96-well plate using *P. torridus* MVD to convert 3-HIV to isobutene. *E. coli* cells engineered with *P. torridus* MVD were grown in a 96-well plate, induced for enzyme expression and then lysed using the previously optimised procedures. Crude extracts were transferred to a fresh plate and mixed with 3-HIV, PEP, ATP, NADH, PK and LDH. Control samples containing empty vector crude extract or only buffer were also analysed. The assay was performed under optimal conditions for activity of PK and LDH, pH 7.5 and 37°C. Activity of *P. torridus* MVD at these conditions was confirmed before performing the assay and

isobutene formation was only 5% slower than maximum rates (247 compared with 261 pmol·min^{-1}·mg$_{(\text{total protein})}^{-1}$).

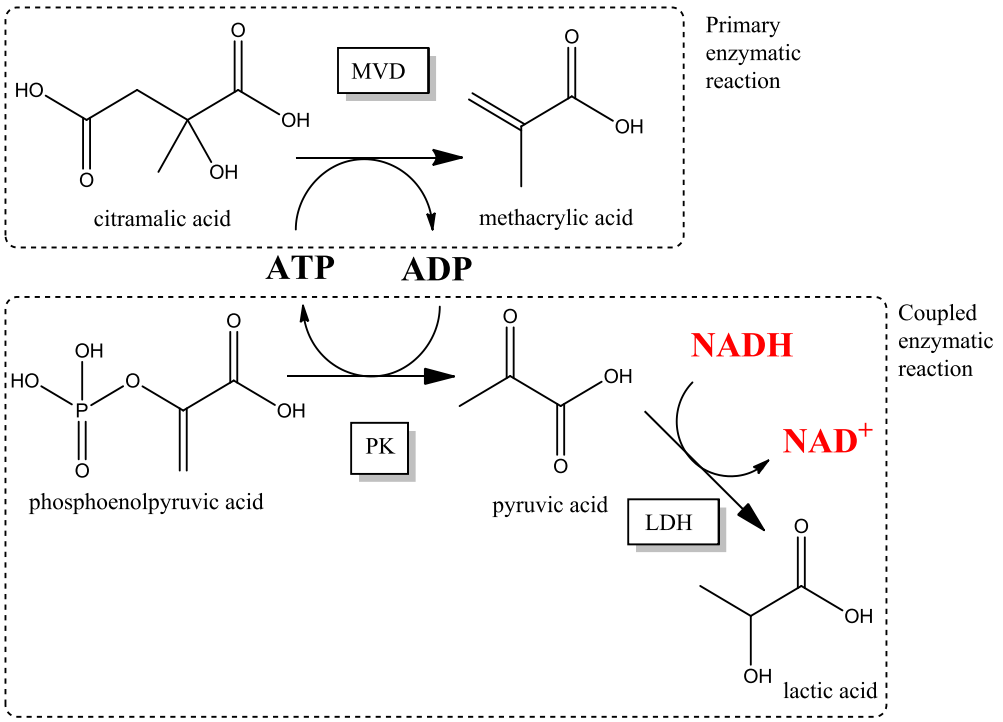


Figure 5.32: Coupled NADH oxidation reaction to monitor ADP release
The enzymatic reaction catalysed by MVD could be monitored by using a coupled enzymatic reaction catalysed by pyruvate kinase (PK) and lactate dehydrogenase (LDH) that oxidises NADH.

Oxidation of NADH was observed in all samples (Table 5.11). In the control samples containing only buffer the rate was very slow, whereas the rate for samples containing empty vector crude extract was very similar to that of the test samples containing MVD. This means that NADH oxidation occurred in cell crude extract, probably due to the presence of NADH oxidases present in *E. coli* or to hydrolysis of ATP. To eliminate the problem, the assay could be done by enzyme purification with Ni-magnetic agarose beads (Hulley *et al.*, 2010) or, since NADH oxidases are usually oxygen dependent, under anaerobic conditions. However, this is laborious, and it was preferred to test other options.

Table 5.11: Evaluation of a NADH oxidation-based screening assay [a]

Only buffer	Empty	MVD
μmol·min^{-1}		
0.102 ± 0.001	0.234 ± 0.003	0.235 ± 0.003

[a] The assay was performed in 96-well plate format (96-well plate) for 48 h. The concentration of 3-HIV used was 5 mM. (Only buffer) refers to samples containing only buffer. (Empty) refers to crude extract from cells containing empty vector. (MVD) refers to samples containing *P. torridus* MVD crude extract. The values indicate the rate of NADH oxidation and are an average of several wells. The means and errors are shown.

5.5.2.1.5 Carbon dioxide release

A last parameter that could be used to monitor MVD activity is the release of CO_2 (Figure 5.23, page 100), perhaps the most obvious approach when screening for decarboxylases. When CO_2 is dissolved in water, a small proportion of it reacts chemically with H_2O to form carbonic acid, H_2CO_3. H_2CO_3 dissociates rapidly to form H^+ ion and HCO^{3-} (bicarbonate ion). This affects the carbonate equilibrium, and the pH value changes as a result (Figure 5.33).

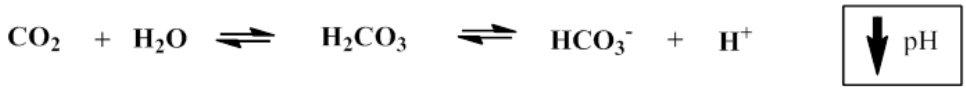


Figure 5.33: pH altered by CO_2in water

Based on this, pH indicators can be used to monitor CO_2 release. They are inexpensive and no auxiliary enzymes are necessary. However, reaction conditions must be chosen carefully to ensure that the colour change is proportional to the number of protons released. In particular, both the buffer and the indicator must have the same affinity for protons (pKas within 0.1 unit of each other) so that the relative amount of protonated buffer to protonated indicator stays constant as the pH shifts during the reaction (Janes *et al.*, 1998). Moreover, no additional protons must be released during the reaction, or this could affect the change in pH as well. Several successful examples employing pH indicators to monitor enzymatic reactions have been reported (Chapman and Wong, 2002; Janes *et al.*, 1998; Rosenberg *et al.*, 1989; Yi *et al.*, 2012).

Therefore, the utilization of a pH indicator as for MVD screening was tested. As before, the assays used *P. torridus* MVD to convert 3-HIV to isobutene in 96-well plate format, in the presence of the pH indicator, bromocresol green. Bromocresol green was used because its pKa is very similar to that of acetate buffer, used for the reaction. The colour change (blue to yellow) was monitored at 612 nm using a plate reader. Controls containing only buffer or empty vector crude were also analysed.

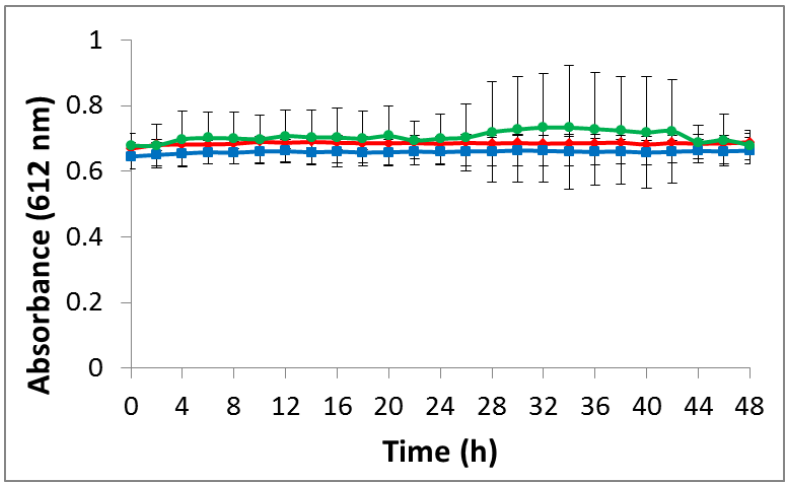


Figure 5.34: Evaluation of a pH indicator-based screening assay
Cell growth, enzyme expression and reaction were all performed in 96-well plate format. Crude extracts were prepared using 50 mM acetate buffer (pH 5.5) with addition of 10 mM $MgCl_2$ and 20 mM KCl. Samples (250 µL) contained crude extract (0.5 mg/mL total protein), 5 mM ATP, 5 mM 3-HIV and 0.05 mM bromocresol green. Plates were incubated for 48 h at 37°C with shaking in a FLUOstar Optima plate reader (BMG Labtech) and changes in bromocresol green absorbance were monitored at 612 nm. The assay was performed with *P. torridus* MVD crude extract (blue line), empty vector crude extract (red line) and only buffer (green line). The curves were obtained as an average of values from several wells.

As expected, no pH change was observed in the controls. However, a pH change was not observed in the test samples either (Figure 5.34). This may have been due to the fact that the release of CO_2 is too low to affect the pH in such a strongly buffered solution, or that the pH decrease due to the release of CO_2 is masked by a pH increase due to the conversion of 3-HIV (an acid) to a neutral gas (isobutene). Sometimes, lowering the buffer strength or increasing the indicator concentration may enhance the sensitivity of the assay (Yi *et al.*, 2012). However, since any activity

on citramalic acid is likely to be very low, it was decided that this screening methodology would not be reliable.

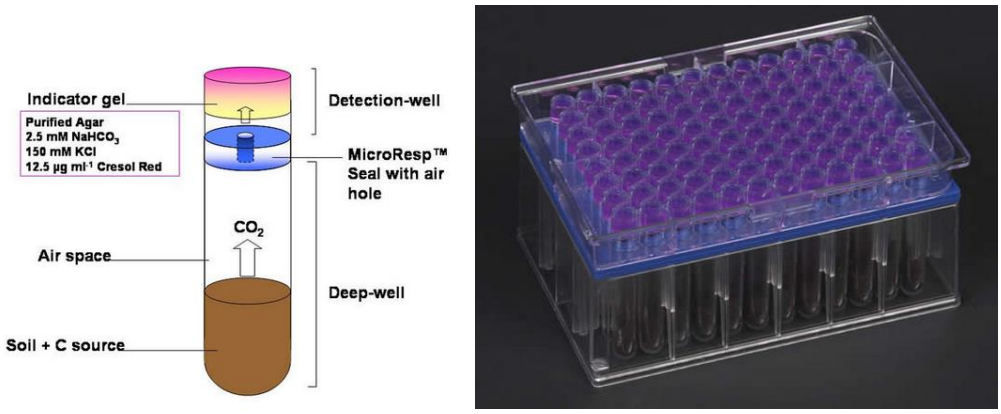


Figure 5.35: Microresp^{TM} method for CO^2 detection in the gas phase
The reaction occurs in a 96-deep well plate. CO_2 is released and interacts with a pH indicator entrapped in agar gel (detection-well). The entrapped pH indicator is prepared in a separate 96-well plate and placed on top of the plate where the reaction occurs. Between the two plates, a PTFE seal avoids leakage of CO_2 into separate wells and a small hole let the gas interact with the detection-well on top of the reaction. Figures are taken from www.microresp.com

Another option for monitoring the release of CO_2 would be to detect it in the gas phase. Microresp^{TM} (Figure 5.35) is a miniaturised method for CO_2 monitoring used in soil respiration analyses (Campbell *et al.*, 2003; Drage *et al.*, 2012; García-Palacios *et al.*, 2013; Lerch *et al.*, 2013; Wakelin *et al.*, 2013), and was therefore investigated. In this system, the CO_2 is released to the headspace of the wells in a 96-deep well plate, where the reaction occurs (Figure 5.35 on the left). The plate is closed with a seal that contains a hole. The hole allows CO_2 to pass through and to interact with a detection well, consisting of bicarbonate and a pH indicator, cresol red, entrapped in agar, previously prepared in another 96-well plate and then assembled upside down on top of the seal (Figure 5.35 on the right). CO_2 reacts with bicarbonate and water, and protons are released ($CO_2^{gas} + H_2O + HCO_3^- \rightarrow 2CO_3^{2-} + 3H^+$). Cresol red changes colour (from pink to yellow) upon release of the protons due to the decrease in the pH. This colour change can be monitored at 570 nm. This methodology has never been applied for

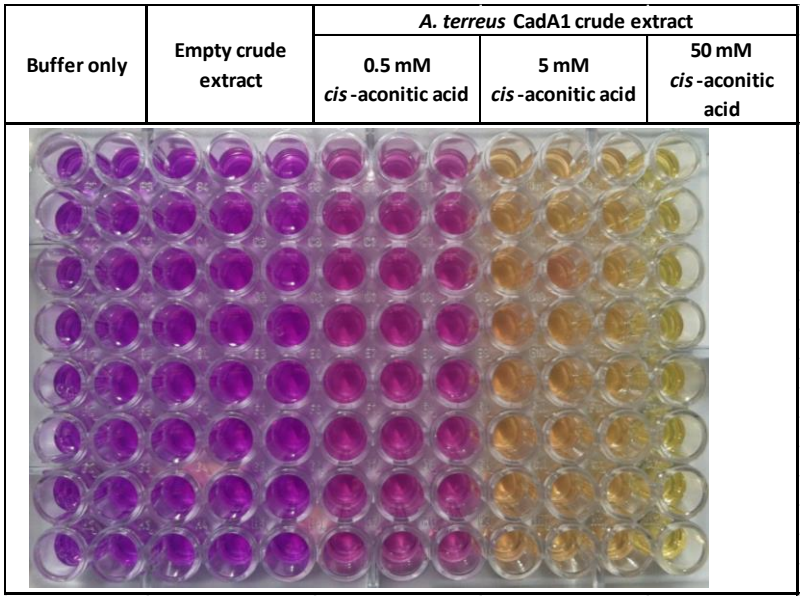
CO_2 monitoring in individual enzymatic reactions, but only with whole cells system from soil extracts (Swallow and Quideau, 2015; Wakelin *et al.*, 2013; Zhou *et al.*, 2013).

The manufacturer suggests a maximum incubation time of 6 h to avoid background decolouration. Therefore, a preliminary evaluation of Microresp™ was first performed using *A. terreus* CadA1, since this enzyme catalysed the decarboxylation of *cis*-aconitic acid to itaconic acid at a very high rate, with complete conversion of 5 mM substrate in less than 5 h (see section 5.2, Page 72). The crude extract of *E. coli* cells engineered with CadA1 was prepared as described in previous chapters and tested on *cis*-aconitic acid in a deep 96-well plate at 30°C. *cis*-Aconitic acid was added at different concentrations (0.5-50 mM) to also test the detection limit of the system. Microresp™ was then assembled and evaluated. Absorbance of detection plates was checked before starting the reaction (A_i) and at the end of the assay (A_f) (3 h). To evaluate CO_2 release, a Decolouring Index (DI) was used, given by the equation $DI=A_f/A_i$. The lower the DI, the higher is the level of CO_2 release. If DI was 1, no decolouration occurred and therefore no CO_2 was released. Controls containing only buffer or empty vector crude extract were also analysed.

A distinct difference in colour change between the controls and the samples containing the active enzyme was clear (Figure 5.36A). Analysis of the DI confirmed this observation (Figure 5.36B). Even at 0.5 mM *cis*-aconitic acid, there was a significant difference in DI from the controls, where no decolouration at all was observed (DI~1). The DI was even lower with higher concentrations of *cis*-aconitic acid. Variations within all samples were minimal. These results proved that Microresp™ can be used to develop a screening assay for decarboxylases, with a very good sensitivity. However, as mentioned above, the manufacturer suggests an incubation time of no more than 6 h to avoid background decolouration. Therefore, longer incubation times were evaluated. Activity of MVD on 3-HIV is slow and may require up to 48 h to develop detectable rates

(Gogerty and Bobik, 2010; Marliere, 2010). Thus, Microresp™ was tested using *P. torridus* MVD and 3-HIV as a substrate (Figure 5.37).

A



B

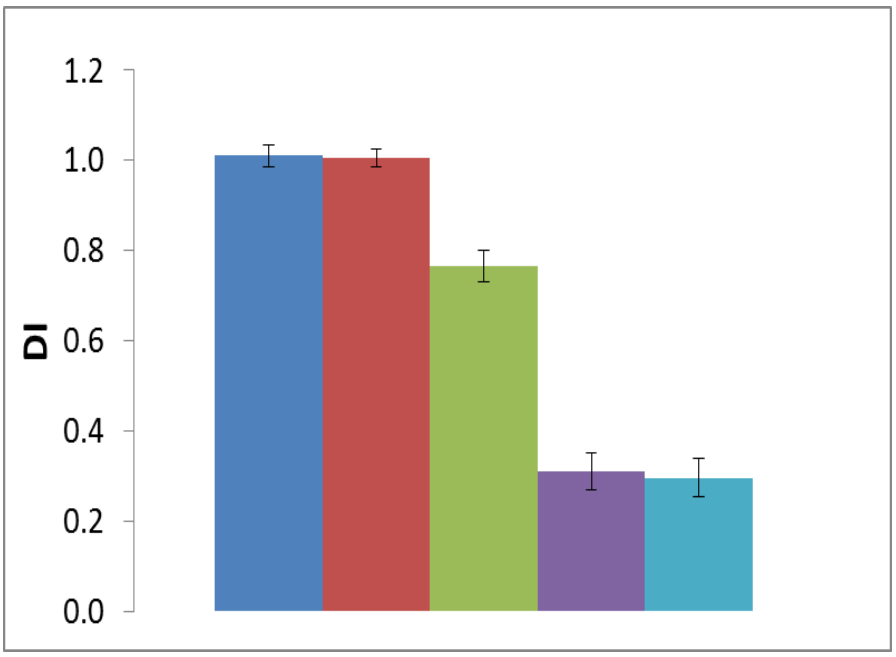
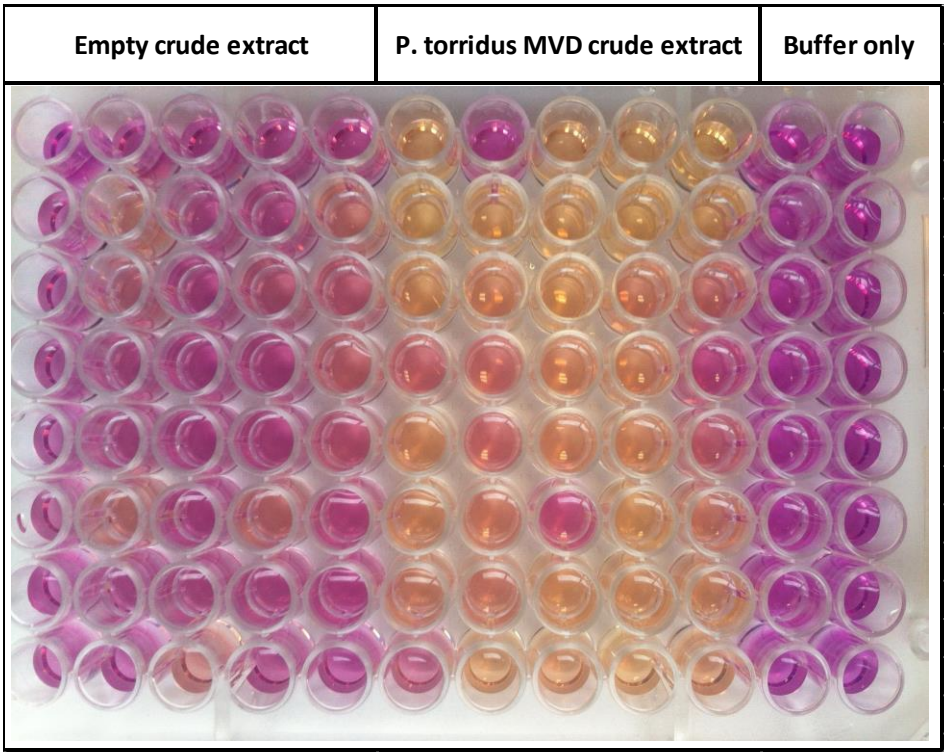


Figure 5.36: Microresp™ evaluation with *A. terreus* CadA1
Detection plate decolouration occurred at the end of the reaction (3 h) (A). Decolouring Index (DI) was calculated after analysing absorbance at 570 nm (B). Control samples contained only buffer (blue bar) or empty vector crude extract (red bar). Test samples contained crude extract from cells expressing *A. terreus* CadA1 and three different concentration of *cis*-aconitic acid: 0.5 (green bar), 5 (purple bar) and 50 (light blue bar) mM. DI values are the average of all replicate wells.

A



B

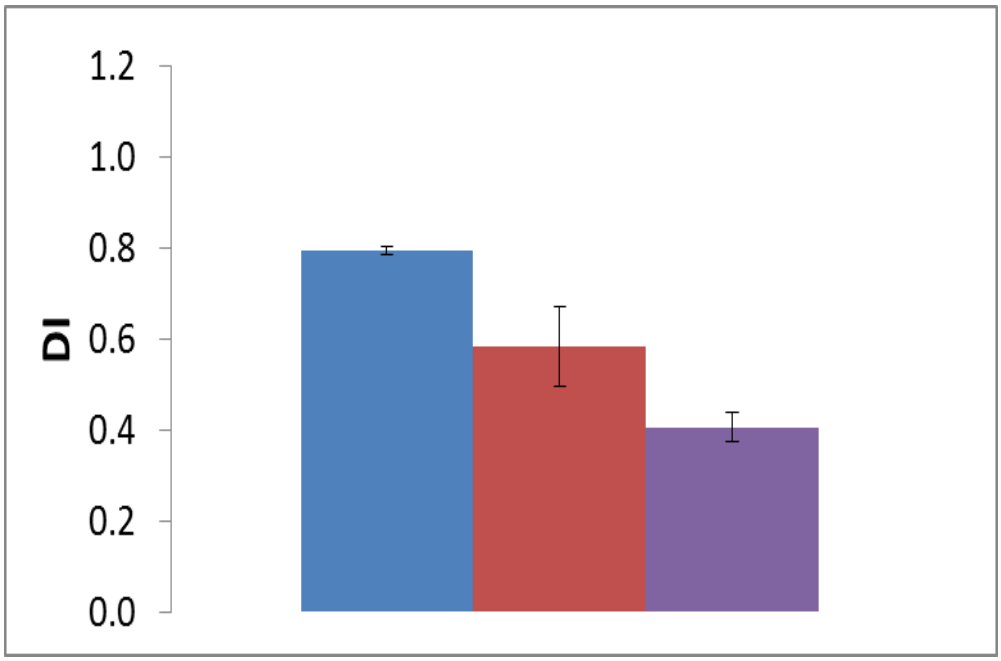


Figure 5.37: Microresp™ evaluation with *P. torridus* MVD
Detection plate decolouration occurred at the end of the reaction (48 h) (A). Decolouring Index (DI) was calculated after analysing absorbance at 570 nm (Bb). Control samples contained only buffer (blue bar) or empty vector crude extract (red bar). Test samples contained crude extract from cells expressing *P. torridus* MVD (purple bar). 3-HIV and ATP were added at a concentration of 5 mM. DI values are the average of all replicate wells.

Cells expressing *P. torridus* MVD were grown, induced and lysed in 96-well plate format. Crude extracts were then transferred to a fresh plate and Microresp™ was assembled. Controls containing only buffer or empty vector crude extract were also analysed. The optimal temperature for *P. torridus* MVD activity, 50°C, could not be employed in the assay in order to avoid evaporation of the reaction mixtures and drying of the detection plates. Therefore, the assay was performed at 37°C. Activity of *P. torridus* MVD at this temperature was confirmed before performing the assay and resulted in only 8% decrease from the maximum rate (240 compared with 261 $pmol \cdot min^{-1} \cdot mg_{(total\ protein)}^{-1}$). The incubation time was prolonged to 48 h.

Compared to the assay performed with *A. terreus* CadA1, even controls containing only buffer showed a minor decolouration (Figure 5.37). Nevertheless, a significant difference between those controls and the test samples containing active MVD indicated that activity of *P. torridus* MVD could be detected even after 48 h. Unexpectedly, control samples containing empty vector crude extract showed a significant decolouration (Figure 5.37) and a high variation, with a coefficient of variance (CV) of 15%. Some wells had a decolouration level significantly higher than others, and this was visible to the naked eye (Figure 5.37A). This was not observed with *A. terreus* CadA1 and was probably due to microbial growth in the reaction mixtures favoured by the long incubation time (Renault *et al.*, 2013) or by un-desired decarboxylation of other metabolic intermediates. The CV in the samples containing active MVD was also rather high, 8%. Nevertheless, the average DI of crude extract controls and of samples containing active MVD was still significantly different. A Z-factor (= 1- [3 x (standard deviation of positive control + standard deviation of negative control)]/ (mean of positive control – mean of negative control)) of 0.015 was calculated, which indicated that the assay could be used (Zhang *et al.*, 1999). However, it is suggested that for a Z-factor very close to 0, the screening assay should be used as "yes or not" type assay rather than a

quantification assay (Zhang *et al.*, 1999). Since, at this stage, a quantitative analysis was not required, this did not present an issue. Taking this statistical analysis into consideration, to minimise the risk of selecting false positives during the screening procedures, it was important to set a cut off value. Ideally, this cut off must be a DI value corresponding to at least 15% less than the average DI of empty vector crude extract controls.

To conclude, among all the tested screening procedures, Microresp™ seemed to be the most reliable. All the other tested screening methodologies showed undesired cross-reactivity with crude extract which limited their employability. Although the results showed that Microresp™ can be slightly affected by a combination of using crude extract and long incubation times, with appropriate precaution (running controls and setting a cut off accordingly), the problem can be minimised. Both activity of *A. terreus* CadA1 and *P. torridus* MVD were, in fact, successfully monitored, even after long incubation times (48 h). The detection limit was excellent. With CadA1, detection was possible even at a CO_2 concentration as low as 0.5 mM and, with *P. torridus* MVD, as low as 0.2 mM (estimated from the activity of the enzyme determined independently). Therefore, in order to proceed with a protein engineering approach to evolve *P. torridus* and *S. cerevisiae* MVD for activity on citramalic acid, Microresp™ was chosen as screening protocol.

A screening assay based on the use of Microresp™ requires several procedure steps. Inevitably, this affects its throughput. Therefore, Microresp™ cannot be used to screen very large numbers of variants, and needs to be combined with a library generation approach that reduces the screening effort. A semi-rational design strategy offers the right option.

5.5.2.2 Library design

5.5.2.2.1 Choice of residues

Identification of target residues is crucial for success in semi-rational design (see section 2.2.2.3, Page 29). In order to select those residues, two criteria were set:

(i) Focus was given to the residues that could affect the structure of the binding pocket;
(ii) Mutagenesis of residues identified as crucial for catalytic activity and cofactor binding was avoided.

Criterion (i) was set because of the difference between citramalic acid and the natural substrate, MVAPP. Whereas the two substrates have the same chemical conformation of the moieties involved in catalysis, citramalic acid presents a carboxylate instead of the classical diphosphate tail of MVAPP. Therefore, modification of the binding pocket to enable the acceptance of the carboxylate was necessary. The crystal structure of *S. cerevisiae* MVD (Bonanno *et al.*, 2001) was used directly to identify those target residues, and (*S*)-citramalic acid was docked *in silico* into the active site to help in the evaluation (Figure 5.38A). For *P. torridus* MVD, no information regarding its structure was available, so residue analysis was done on a homology model, obtained using I-TASSER (Roy *et al.*, 2010; Zhang, 2008) (Figure 5.38B). The model was considered reliable based on the excellent C-score (+1.43 on a scale of -5 to +2), a confidence score provided by the I-TASSER server for estimating the quality of the predicted model. (*S*)-citramalic acid could not be docked into the predicted active site as for *S. cerevisiae* MVD since attempts to obtain valid docking results were poor. Therefore, the model was first superimposed on *Staphylococcus epidermidis* MVD (Barta *et al.*, 2012; Barta *et al.*, 2011), and citramalic acid was built on the scaffold of the co-crystallised MVAPP.

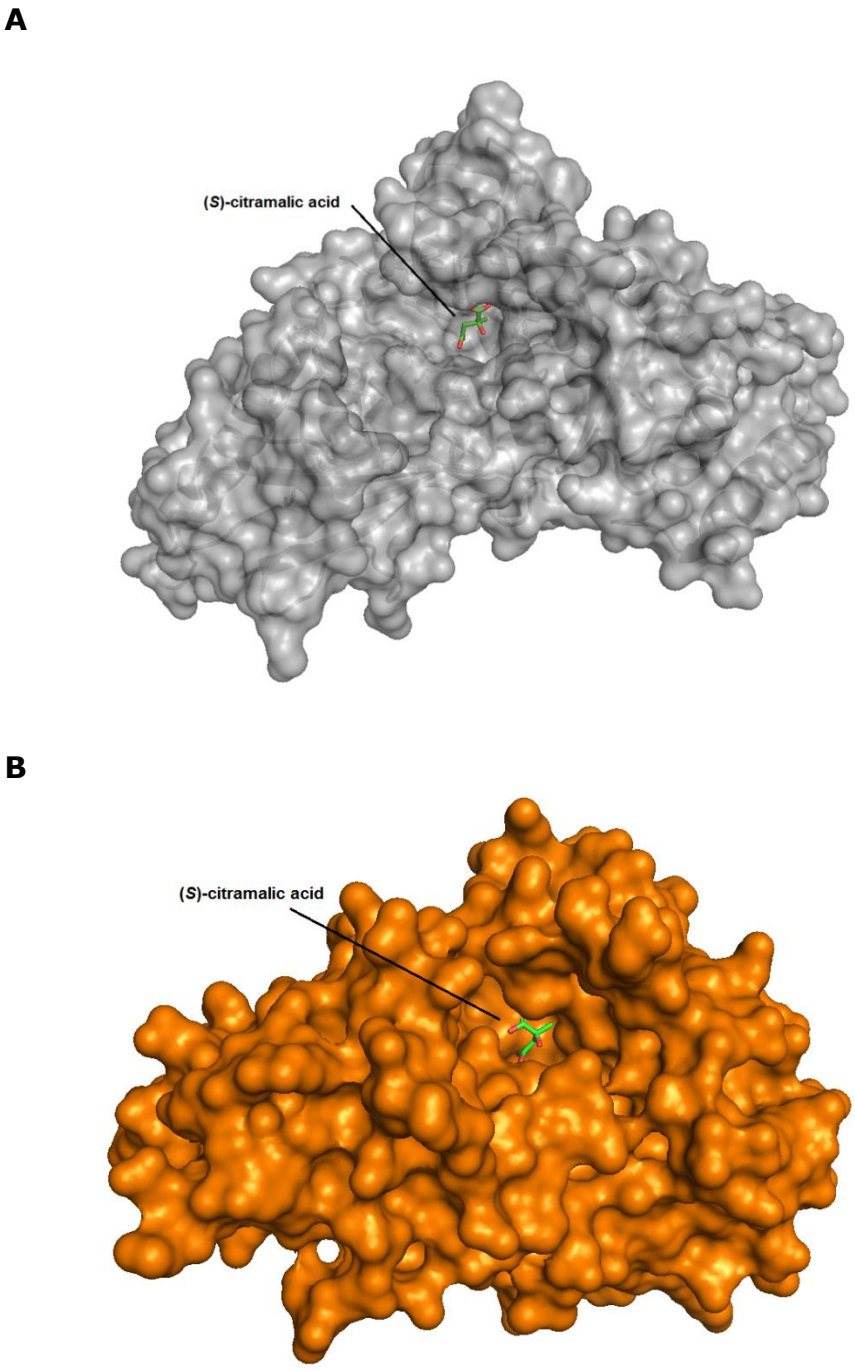


Figure 5.38: Structure of *S. cerevisiae* and *P. torridus* MVD
The crystal structure *of S. cerevisiae* MVD (A) was used directly for residue analysis and (*S*)-citramalic acid was docket *in silico* into the active site. For *P. torridus* MVD (B), a homology model was instead used and (*S*)-citramalic acid was built into the active site on the scaffold of the natural substrate of other MVDs used as template for the model.

Activity of MVD involves several residues, some directly participating in catalysis, others required for binding ATP and for structural rearrangements of the enzyme. Criterion (ii) was set to avoid touching these residues and introducing counterproductive

mutations that could kill enzymatic activity. To gain more insights into the function of the residues around the regions of interest, I gathered all available information from any mutagenesis, inhibition, structural and mechanistic studies on MVDs (Alvear *et al.*, 1982; Barta *et al.*, 2012; Barta *et al.*, 2011; Bonanno *et al.*, 2001; Byres *et al.*, 2005; Cordier *et al.*, 1999a; Cordier *et al.*, 1999b; Dassanayake R. *et al.*, 2002; Krepkiy and Miziorko, 2004; Michihara *et al.*, 2002; Pang *et al.*, 2006; Qiu *et al.*, 2007; Qiu and Li, 2006; Shi *et al.*, 2011; Voynova *et al.*, 2008; Weerasinghe and Dassanayake, 2009) (Table 5.12). Subsequently, I related the information regarding the function of specific residues to the corresponding residues of *P. torridus* and *S. cerevisiae* MVDs by doing a multiple sequence alignment (MSA) for a cohort of different MVDs (Figure 5.39).

Table 5.12: Summary of information regarding residues important for MVD function

Residue ***S. epidermidis*** **numbering scheme**	**Suggested function**	***S. cerevisiae*** **MVD corresponding residue**	***P. torridus*** **MVD corresponding residue**	**Ref.**
Asn12	- Helps in positioning the residues that interacts with C1 of MVAPP	Asn13	Thr13	(Voynova *et al.*, 2008)
Lys17	- Binds MVAPP - Interacts with some inhibitor	Lys18	Leu18	(Barta *et al.*, 2012; Barta *et al.*, 2011; Krepkiy and Miziorko, 2004; Weerasinghe and Dassanayake, 2009)
Tyr18	- May interact with the phosphates of ATP or MVAPP - Interacts with some inhibitor - Helps in orienting catalytic Arg	Tyr19	Leu 19	(Barta *et al.*, 2012; Barta *et al.*, 2011; Weerasinghe and Dassanayake, 2009)
Trp19	- Helps in forming an hydrophobic region to allow binding pocket folding	Trp20	Gly 20	(Byres *et al.*, 2007)
Lys21	- Binds the phosphates of MVAPP	Lys22	Ile22	(Barta *et al.*, 2012; Barta *et al.*, 2011; Byres *et al.*, 2007; Weerasinghe and Dassanayake, 2009)
Asn61	- Interacts with the adenine moiety of ATP	Asn64	Asp63	(Byres *et al.*, 2007)
Glu69	- Interacts with the ribose moiety of ATP	Asn72	Ser71	(Barta *et al.*, 2012; Barta *et al.*, 2011)
Lys72	- Binds the phosphates of MVAPP	Arg74	Ser74	(Barta *et al.*, 2012; Barta *et al.*, 2011; Byres *et al.*, 2007; Gogerty and Bobik, 2010)

Ser94	- Interacts with the adenine moiety of ATP	Ser108	Asp101	(Barta *et al.*, 2012; Barta *et al.*, 2011)
Asn96	- Interacts with the adenine moiety of ATP	Asn110	Arg103	(Barta *et al.*, 2012; Barta *et al.*, 2011; Byres *et al.*, 2007)
Ala105	- Interacts with the phosphates of ATP	Ala119	n/a	(Barta *et al.*, 2012; Barta *et al.*, 2011)
Ser106	- Interacts with the phosphates of ATP	Ser120	Ser111	(Barta *et al.*, 2012; Barta *et al.*, 2011; Byres *et al.*, 2007; Krepkiy and Miziorko, 2005)
Ser107	- Interacts with the phosphates of ATP - Helps in correctly orienting MVAPP in the active site	Ser121	Ser112	(Barta *et al.*, 2012; Barta *et al.*, 2011; Byres *et al.*, 2007; Krepkiy and Miziorko, 2005)
Ala108	- Interacts with the phosphates of ATP	Ala122	Asp113	(Barta *et al.*, 2012; Barta *et al.*, 2011)
Tyr111	- Interacts with the adenine moiety of ATP	Phe125	Ala116	(Byres *et al.*, 2007)
Ser139	- Binds the phosphates of ATP and MVAPP	Ser153	Ser154	(Barta *et al.*, 2012; Barta *et al.*, 2011; Byres *et al.*, 2007; Krepkiy and Miziorko, 2004; Krepkiy and Miziorko, 2005; Weerasinghe and Dassanayake, 2009)
Gly140	- Binds the phosphates of MVAPP - May interact with the phosphates of ATP	Gly154	Asp155	(Barta *et al.*, 2012; Barta *et al.*, 2011; Byres *et al.*, 2007; Krepkiy and Miziorko, 2004; Krepkiy and Miziorko, 2005; Weerasinghe and Dassanayake, 2009)
Ser141	- Binds the phosphates of ATP and MVAPP	Ser155	Ser156	(Barta *et al.*, 2012; Barta *et al.*, 2011; Byres *et al.*, 2007; Krepkiy and Miziorko, 2004; Krepkiy and Miziorko, 2005; Weerasinghe and Dassanayake, 2009)
Arg144	- Binds C1 of MVAPP and drives decarboxylation	Arg158	Arg149	(Barta *et al.*, 2012; Barta *et al.*, 2011; Byres *et al.*, 2007; Qiu *et al.*, 2007; Voynova *et al.*, 2008)
Trp153	- Helps in forming an hydrophobic region to allow binding pocket folding	Trp167	Tyr159	(Byres *et al.*, 2007)
Ser192	- Binds the phosphates of ATP and MVAPP and help their positioning	Ser208	Glu195	(Barta *et al.*, 2012; Barta *et al.*, 2011)
Arg193	- Binds the phosphates of MVAPP	Thr209	Ile196	(Barta *et al.*, 2012; Barta *et al.*, 2011; Byres *et al.*, 2007; Weerasinghe and Dassanayake, 2009)
Met196	- Helps orienting MVAPP in the active site	Met212	Asn199	(Byres *et al.*, 2007)
Asp283	- Directly involved in catalysis - Deprotonates MVAPP C3 to favour phosphorylation	Asp302	Thr280	(Barta *et al.*, 2012; Barta *et al.*, 2011; Byres *et al.*, 2007; Qiu *et al.*, 2007; Voynova *et al.*, 2008)
Ala284	- May help in orienting catalytic Asp	Ala303	Gly281	(Barta *et al.*, 2012; Barta *et al.*, 2011)

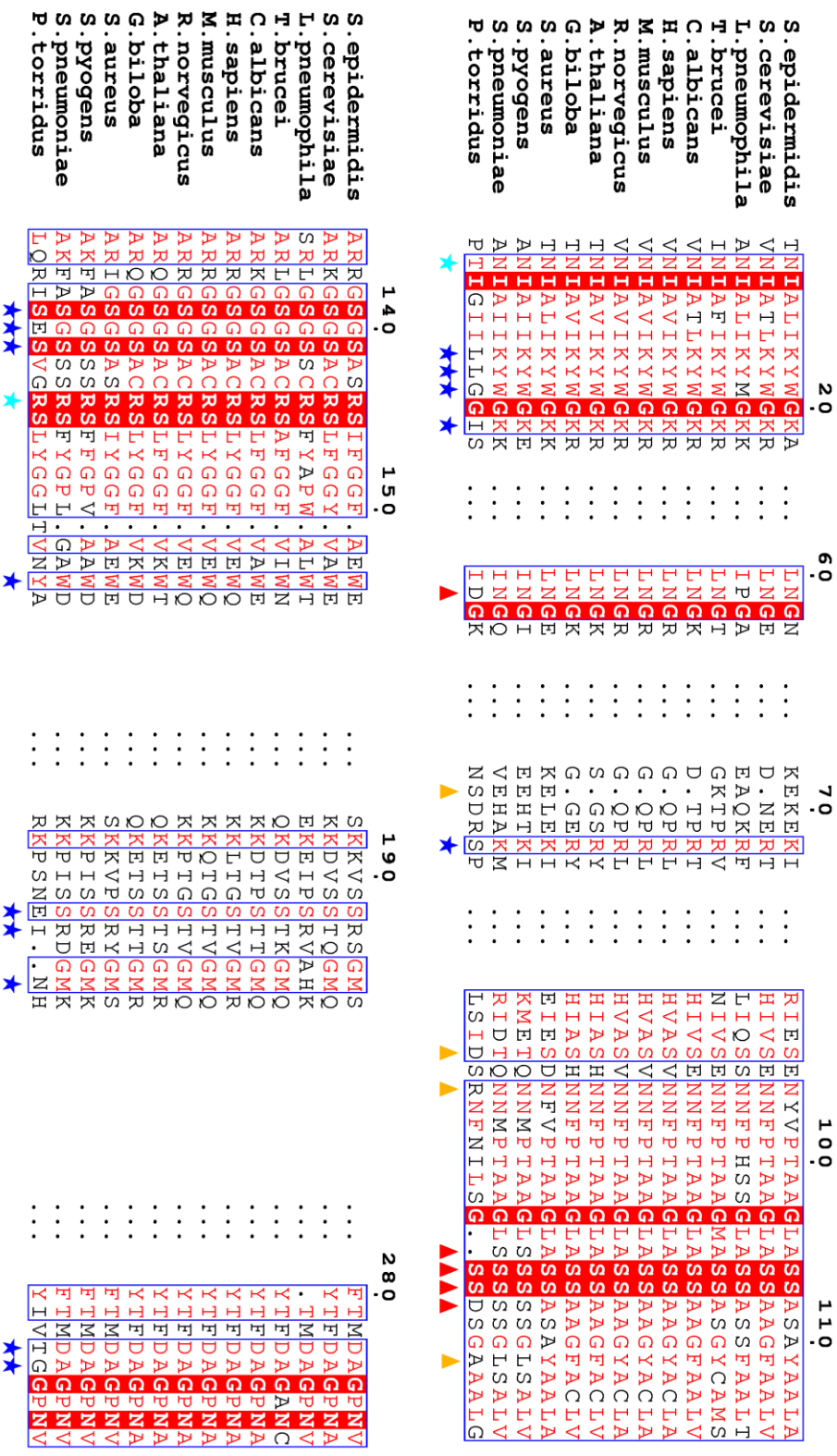


Figure 5.39: Multiple sequence alignment of MVD homologues.
Highly conserved regions are framed in blue. Conservation of residues is denoted by colours: white amino acids highlighted in red are strictly conserved; red amino acids not highlighted are high in similarity. The numbering annotation is based on *S. epidermidis* MVD. Stars denote residues involved in MVAPP interaction (cyan for C1 carboxylate group, blue for diphosphate moiety and hydroxyl group). Triangles denote residues involved in ATP interaction (red for phosphate groups, orange for purine ring and ribose).

For *P. torridus* MVD, several residues around the binding pocket were not considered for mutagenesis because of criterion (ii) (example in Figure 5.40A). Others were not considered because, although within the binding pocket, they were located in a position that was unlikely to affect acceptance of the extra carboxylic moiety of citramalic acid (example in Figure 5.40B). Also for *S. cerevisiae* MVD, several residues were not considered for the same reasons. In the end, only four amino acids could be chosen for each enzyme: Gly20, Ile22, Ile196 and Asn199 (Figure 5.41) for *P. torridus* MVD; Gly154, Ser155, Ser208 and Met212 (Figure 5.42) for *S. cerevisiae* MVD.

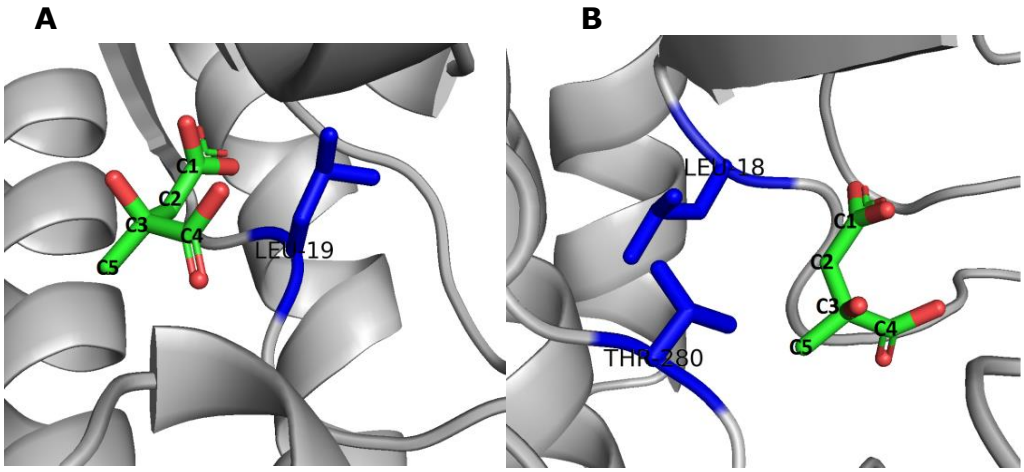


Figure 5.40: Residues evaluation for library design of *P. torridus* MVD (Leu18, Leu19, Thr280)
Residues are highlighted in blue and citramalic acid is depicted in green (carbon atoms) and red (oxygen atoms). Carbons 1-5 of citramalic acid are labelled. Leu19 (A), although in good position to modify the binding pocket around citramalic acid's C3 carboxylate moiety, was suggested to be involved in orienting of Arg148, crucial for activity (Byres *et al.*, 2007). Leu18 and Thr280 (B), are located near the substrate's C1 rather than the C3 carboxylate moiety.

Whereas for *P. torridus* MVD, the selected residues respected both criteria (i) and (ii), for *S. cerevisiae* MVD this was not possible. Three amino acids among those selected (Gly154, Ser155 and Ser208) are in fact believed to participate in the coordination of ATP (Barta *et al.*, 2012; Barta *et al.*, 2011; Byres *et al.*, 2007; Krepkiy and Miziorko, 2004; Krepkiy and Miziorko, 2005; Weerasinghe and Dassanayake, 2009), and this is against criterion (ii). Nevertheless, these residues were the best option, as there was no other residue that could respect criterion (ii) and be in a good location for the

binding of citramalic acid at the same time. Moreover, since these residues are suggested to only partially coordinate ATP, they were less likely than others to affect the enzymatic mechanism

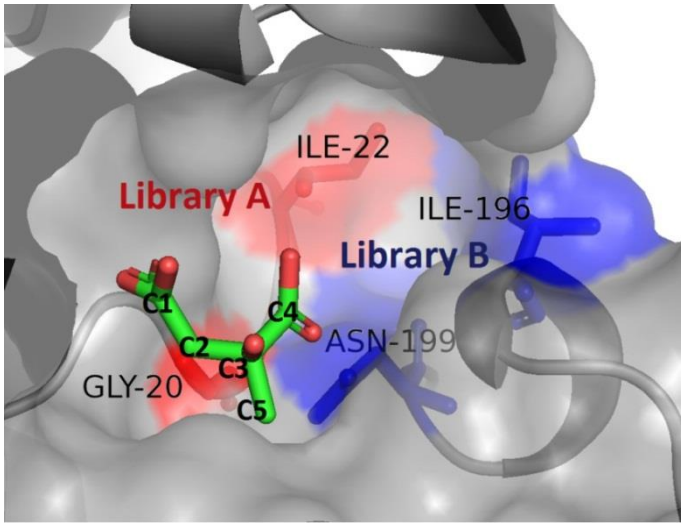


Figure 5.41: Residues selected for mutagenesis of *P. torridus* MVD
Residues highlighted in red were mutated in Library A. Residues highlighted in blue were mutated for Library B. Citramalic acid is depicted in green (carbon atoms) and red (oxygen atoms). Carbons 1-5 of citramalic acid are labelled.

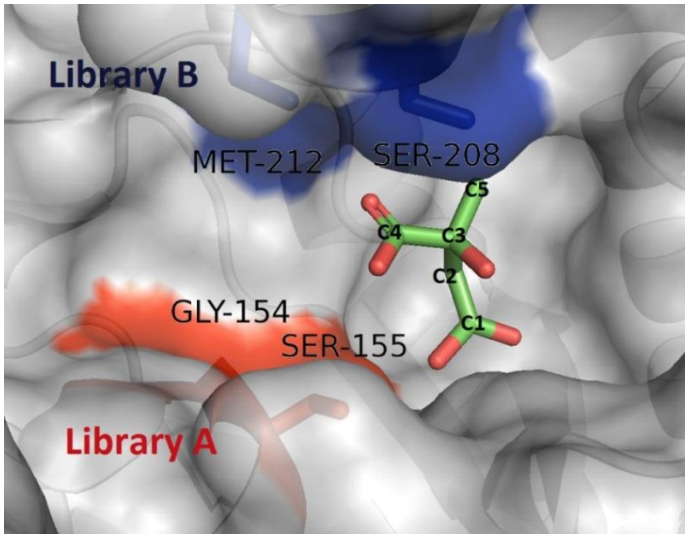


Figure 5.42: Residues selected for mutagenesis of *S. cerevisiae*
Residues highlighted in red were mutated in Library A. Residues highlighted in blue were mutated for Library B. Citramalic acid is depicted in green (carbon atoms) and red (oxygen atoms). Carbons 1-5 of citramalic acid are labelled.

5.5.2.2.2 Introduction of variation

In order to choose how to introduce variation at the decided positions, the following criteria were set:

(a) The corresponding wild-type amino acid at each position was included in the amino acid pool introduced;

(b) A structurally balanced mixture of amino acids (with polar, non-polar, aromatic, and lipophilic side chains) was preferred;
(c) Amino acids which were more likely to be potentially beneficial for citramalic acid acceptance were included as much as possible;
(d) The number of variants to be screened for each enzyme could not exceed 2000 for each enzyme (4000 in total);

Criterion (a) was necessary to assure that one among the amino acid could remain non-mutated. Criteria (b) and (c) were set to maximise the number of potential modifications in the binding pocket of the enzyme and the likelihood of success. Criterion (d) was set so that Microresp^{TM} could be used as the screening method in reasonable time scale.

To respect all the criteria, the right balance between the number of positions that could be simultaneously mutated and the level of degeneracy to introduce during mutagenesis needed to be found. The higher the number of residues and the degree of codon degeneracy, the more variants would need to be screened. Moreover, for a certain number of variants generated, about threefold excess samples need to be analysed to statistically cover a reliable number of the possible combinations of mutations, referred to as oversampling (Bosley and Ostermeier, 2005). This further affects the screening effort required. Therefore, the libraries had to be designed so that the oversampling required was low, but also in a way in which the variation introduced could still be in line with criteria (a), (b) and (c). The solution to these problems was to follow the principles of ISM (Reetz and Carballeira, 2007), a focused mutagenesis approach that splits the residues chosen for mutagenesis into separate, smaller libraries, and that subsequently uses random mutagenesis iteratively at each library until the desired grade of enhancement is reached (see section 2.2.2.3, Page 29). ISM has the capacity to reduce screening effort significantly without

compromising, and rather increasing, the likelihood of evolving enhanced enzymes (Reetz and Sanchis, 2008).

The four residues selected for both *S. cerevisiae* and *P. torridus* MVD were split into two separate groups of only two amino acids. Simultaneous mutagenesis at these positions was then performed to generate two different libraries for each enzyme (A and B). For *P. torridus* MVD, Library A was generated by simultaneous mutagenesis of Gly20 and Ile22, and Library B of Ile196 and Asn199 (Figure 5.41, Page 129). For *S. cerevisiae* MVD, Library A was generated by simultaneous mutagenesis of Gly154 and Ser155, and Library B of Ser208 and Met212 (Figure 5.42, Page 129). Residues in close proximity were used for the same library to stimulate synergistic effects (Reetz *et al.*, 2010a). The plan was then to select the best enzyme variants generated for each library and iteratively perform other rounds of mutagenesis at the second site, according to the principles of ISM.

CASTER, a computer program that aids in designing saturation mutagenesis libraries (Reetz and Carballeira, 2007), was used to decide the codon degeneracy to use at each position and to evaluate the screening effort required. NDT (Reetz, 2011; Reetz *et al.*, 2008) was initially taken into consideration. With simultaneous mutagenesis of two residues for each library, the number of clones to be screened using NDT would have been 860 for each enzyme, well below the limit set in criterion (d). However, NDT does not encode Trp and Lys, two residues which were considered important. Trp, a bulky amino acid, could help in shortening the binding pocket and develop acceptance of a short substrate such as citramalic acid. Lys, positively charged, may form an interaction with the negatively charged carboxylate of citramalic acid. Therefore, another codon degeneracy was taken into consideration, NDK, which encodes for 17 amino acids including Trp and Lys. However, by using NDK degeneracy at two positions simultaneously, the number of variants to be screened would have been 3448 for each enzyme (~7000 in total), beyond the limit set. Therefore, it was decided to use a

combination of NDT and NDK codon degeneracies, where NDT was used for one position of each library, and NDK for the other one. With this adjustment, the number of clones to be screened for each enzyme was lowered to 1722.

For *P. torridus* MVD, the NDK codon degeneracy was used for the residues closer to citramalic acid (Gly20 and Asn199) (Table 5.13), since they were more likely to interact with the substrate. The NDT codon degeneracy was instead used for the residues that are deeper in the binding pocket (Ile22 and Ile196) (Table 5.13). For *S. cerevisiae* MVD, since both residues of each library are almost at equal distances from the carboxylate moiety of citramalic acid, the decision on what degeneracy to use at each position was arbitrary. The NDT codon degeneracy was chosen for Ser155 and Ser208 and the NDK codon degeneracy for mutagenesis at Gly155 and Met212 (Table 5.14)

Table 5.13: Library decision summary for *P. torridus* MVD

Library	Position	Codon	Encoded amino acids
A	Gly20	NDK	Arg, Asn, Asp, Cys, Gln, Glu, Gly, His, Ile, Leu, Lys, Met, Phe, Ser, Trp, Tyr, Val
	Ile22	NDT	Arg, Asn, Asp, Cys, Gly, His, Ile, Leu, Phe, Ser, Tyr, Val
B	Ile196	NDT	Arg, Asn, Asp, Cys, Gly, His, Ile, Leu, Phe, Ser, Tyr, Val
	Asn199	NDK	Arg, Asn, Asp, Cys, Gln, Glu, Gly, His, Ile, Leu, Lys, Met, Phe, Ser, Trp, Tyr, Val

Table 5.14: Library decision summary for *S. cerevisiae* MVD

Library	Position	Codon	Encoded amino acids
A	Gly154	NDK	Arg, Asn, Asp, Cys, Gln, Glu, Gly, His, Ile, Leu, Lys, Met, Phe, Ser, Trp, Tyr, Val
	Ser155	NDT	Arg, Asn, Asp, Cys, Gly, His, Ile, Leu, Phe, Ser, Tyr, Val
B	Ser208	NDT	Arg, Asn, Asp, Cys, Gly, His, Ile, Leu, Phe, Ser, Tyr, Val
	Met212	NDK	Arg, Asn, Asp, Cys, Gln, Glu, Gly, His, Ile, Leu, Lys, Met, Phe, Ser, Trp, Tyr, Val

5.5.2.3 Screening results

5.5.2.3.1 Library generation

Libraries were generated using the QuickChange Lightening Site-Directed Mutagenesis Kit (Agilent) (Figure 5.43) which is based on the principle of whole plasmid amplification with mutagenic primers (Papworth *et al.*, 1996). pET-20b(+) plasmids containing the genes encoding *P. torridus* MVD and *S. cerevisiae* MVD were used as templates. For each library, mutations were in close proximity. Therefore, it was possible to design mutagenic primers carrying both codon degeneracies at the same time (Table 4.2, Page 44).

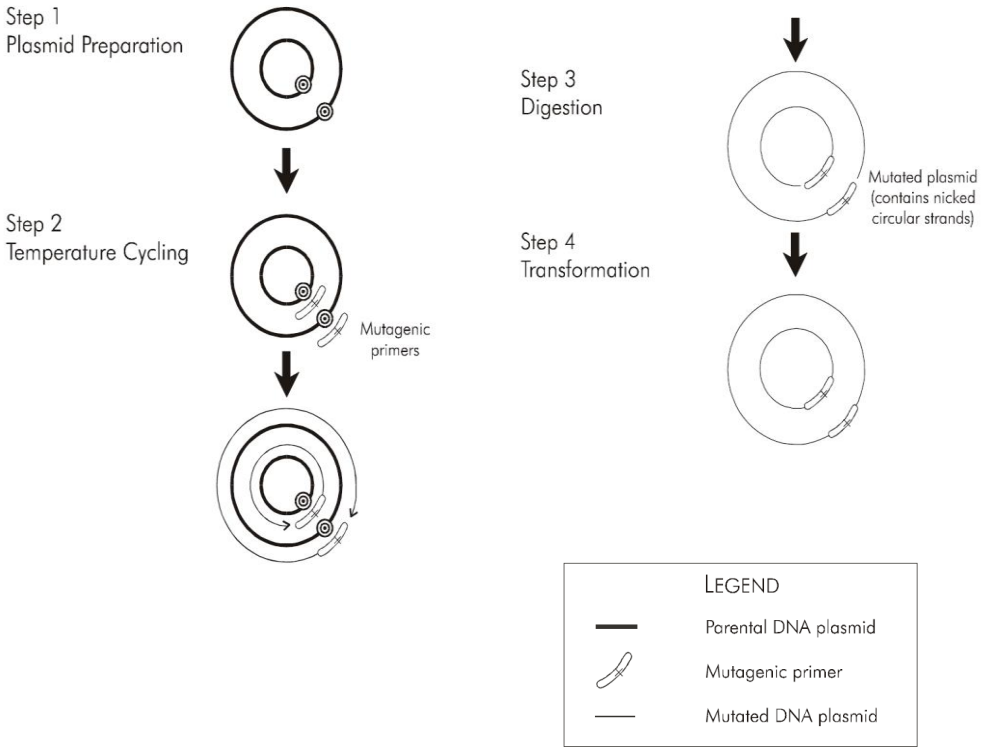


Figure 5.43: Overview of the QuickChange mutagenesis method
Step 1: a methylated double-stranded plasmid containing the target wild type gene is obtained by purification from an appropriate *E. coli* strain. Step2: two mutagenic primers containing the desired DNA modifications are designed to be complementary to the same region of the vector on opposite strands. The vector is then amplified by PCR using those primers. The resulting amplification product is a non-methylated plasmid library presenting the same modifications contained in the primers. Since the DNA polymerase used in the reaction is not able to re-connect separated DNA fragments, the plasmids obtained by amplification are characterised by the presence of staggered nicks. Step 3: to eliminate parental plasmid containing the non-mutated gene, the amplification product is treated with *Dpn I,* an endonuclease that specifically acts on methylated DNA. Step 4: the plasmid library is transformed into competent *E. coli* cells which are able to repair the nicked vectors. Plasmid libraries can then be amplified, purified and frozen, ready to be used and screened after expression of the enzyme in an expression host**.** The figure is taken from the manufacturer's manual.

Amplification was performed using different concentrations of template plasmids and, after removal of parental DNA through *Dpn I* treatment, *E. coli* XL-Gold Ultracompetent cells were transformed and plated on agar. The reason for testing different concentrations of template plasmids was to optimise the number of clones generated, aiming to obtain a number which exceeded the oversampling requirement (861 for each library). This precaution made sure that the generated libraries were big enough to include the highest possible number of mutant combinations. For *P. torridus* MVD, the optimal concentration of template plasmid was 1 ng/µL, whereas for *S. cerevisiae* MVD was 1.5 ng/µL. In both cases, more than 3000 clones were generated.

To confirm that the libraries contained the expected degeneracy at the decided positions, the transformed cells were also grown in fresh medium and the generated libraries were extracted. Plasmids were then sent for sequencing and the resulting chromatograms analysed. In a sequencing process based on the Sanger method (as used by the company doing the analysis), a DNA sequence is analysed using fluorescent nucleobases (Sanger and Coulson, 1975). The output of the analysis is a chromatogram in which each nucleobase appears as a peak of a different colour. When multiple peaks are visible at a certain position, it means that the DNA sequence analysed contains a mixture of different nucleobases. For each codon degeneracy (NDK and NDT), the respective mixture of nucleobases was expected. N should have given a mixture of all A-T-C-G, D, a mixture of A-G-T, and K, a mixture of T-G. As visible from selected sections of the chromatograms (Figure 5.44), the mixture of nucleobases found at each position was as expected. Therefore, the purified plasmid libraries were used for screening. As the area of the peaks was similar for all the relative nucleobases expected at each position, indicating minimal codon bias, sequence analysis of individual clones was not performed.

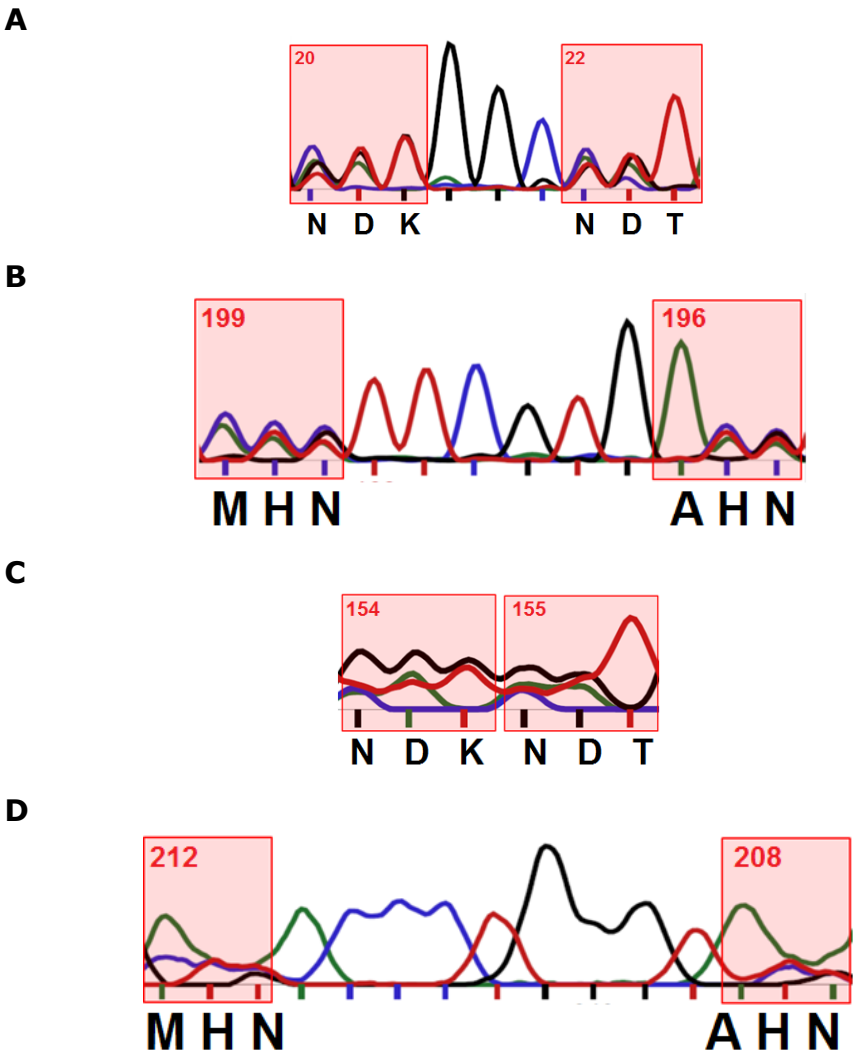


Figure 5.44: Sequencing results chromatograms of MVD plasmid libraries
Only the chromatogram sections referring to the positions where mutagenesis was performed are shown. Mutagenic codons are highlighted in red boxes. For each codon, the position number of the respective encoded amino acid in *P. torridus* and *S. cerevisiae* MVD is indicated in red. (A) and (B) refer to the chromatograms obtained from the analysis of library A and B generated from the plasmid containing the *P. torridus* MVD gene. (C) and (D) refer to the chromatograms obtained from the analysis of library A and B generated from the plasmid containing the *S. cerevisiae* MVD gene. The expected codon degeneracy is indicated below the red boxes. Blue peaks correspond to C, black peaks correspond to G, green peaks correspond to A and red peaks correspond to T. Some of the sequencing analyses of the double-stranded plasmids were done on one strand, (A) and (C), and some were done on the complementary strand, (B) and (D). In the latter cases, the chromatograms show the MHN and AHN codon degeneracies, which are complementary to the NDK and NDT, respectively.

5.5.2.3.2 Microresp™ screening

The next step was to transform an expression host with the successfully generated libraries, and to screen enzyme variants for MAA formation, indicated by CO_2 release from

dehydration/decarboxylation of citramalic acid, using Microresp™. *E. coli* BL21(DE3)pLysS was transformed and plated on agar at a dilution that allowed growth of at least 861 colonies, the minimum required for a good oversampling. Single colonies were then individually grown, induced and lysed in 96-well plate format. Crude lysates were transferred to a fresh 96-well plate where (*S*)-citramalic acid (10 mM) and ATP (10 mM) were added to start the reaction. For each plate, the crude extract randomly selected from 8-10 mutants was analysed by SDS-PAGE to evaluate enzyme expression and the total protein concentration. Microresp™ was finally assembled over the plates and incubated at 37°C for 48 h. When testing variants of *P. torridus* MVD, reactions were performed at pH 5.5, as these conditions were already successful during Microresp™ optimisation. When testing variants of *S. cerevisiae* MVD, reactions were performed at pH 7, although the optimal pH for enzyme activity is 7.5 (Gogerty and Bobik, 2010). The pH was lowered because Microresp™ requires a pH $\leq$ 7 to maximise sensitivity (Renault *et al.*, 2013). Activity of *S. cerevisiae* MVD at this pH was confirmed before performing the assay.

Absorbance of the detection plates of Microresp™ was analysed before assembling the apparatus (A_i) and at the end of the reaction (A_f). The evaluation of the DI ($=A_f/A_i$) for each variant was used to understand whether any release of CO_2 occurred. To minimise the possibility of false positives, a cut off DI was set for each individual plate screened. Negative controls containing empty vector crude extract were analysed alongside the test samples and the cut off was set as the value corresponding to 16% lower than their average DI, as suggested from the results obtained during the optimization of the assay.

A total of 3840 clones, 1920 for each enzyme (960 for each library), were analysed and 48 Microresp™ assays were necessary (12 plates for each library). Several positive hits were observed, almost evenly distributed among all plates. For *P. torridus* MVD Library A and B, 27 and 22 positive hits were obtained, respectively

(Figure 5.45), whereas for *S. cerevisiae* MVD the number was higher, 45 for Library A and 50 for Library B (Figure 5.46). Among these hits, some of the DIs were only slightly lower than the cut off and some were instead characterised by a more significant difference, also visible to the naked eye (Figure 5.47).

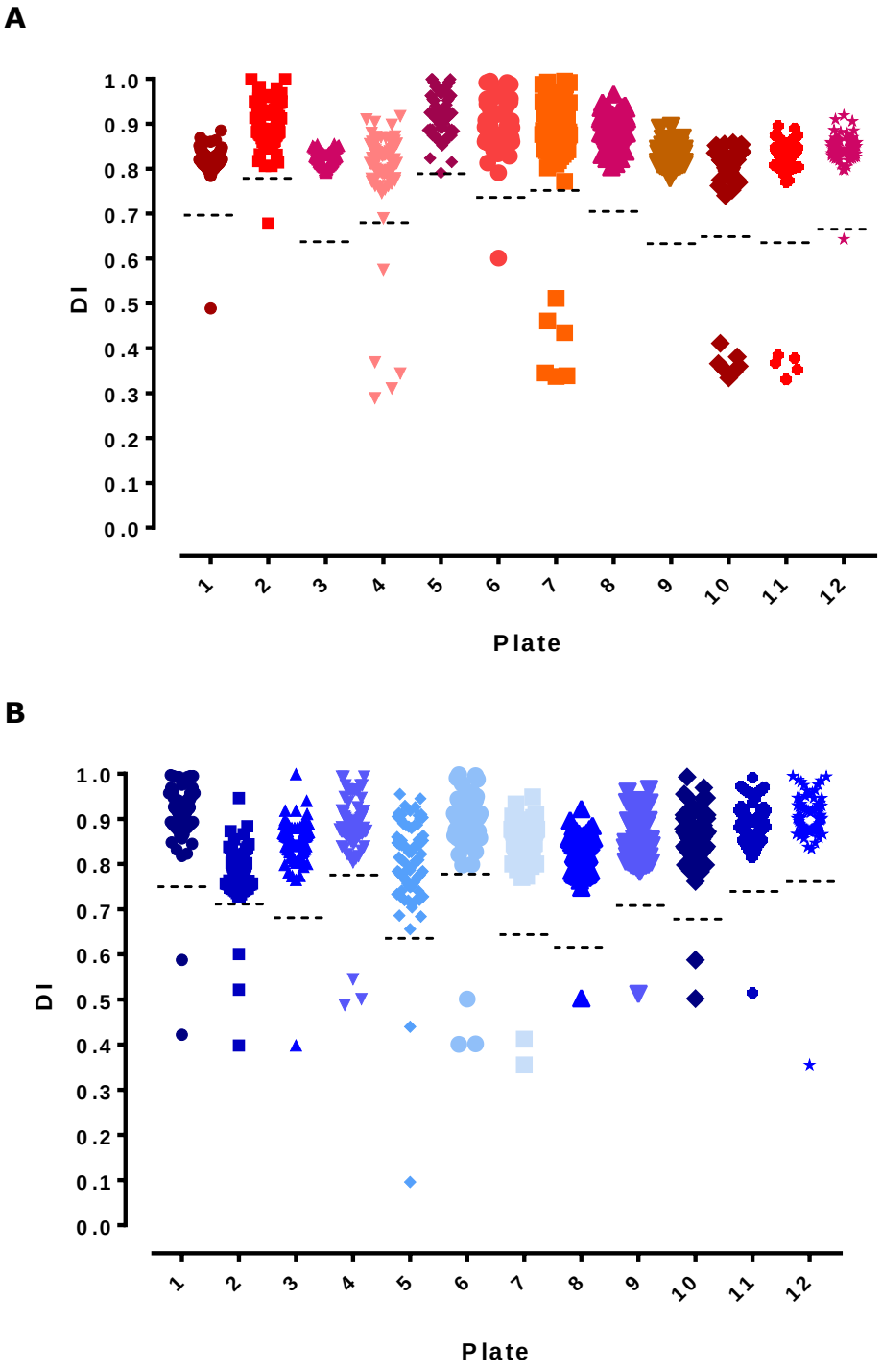


Figure 5.45: Screening results for *P. torridus* MVD variants
(A) refers to Library A and (B) to Library B. DI values for the variants are indicated for each plate screened (1-12). The cut off limit set is indicated for each plate with a dashed black line.

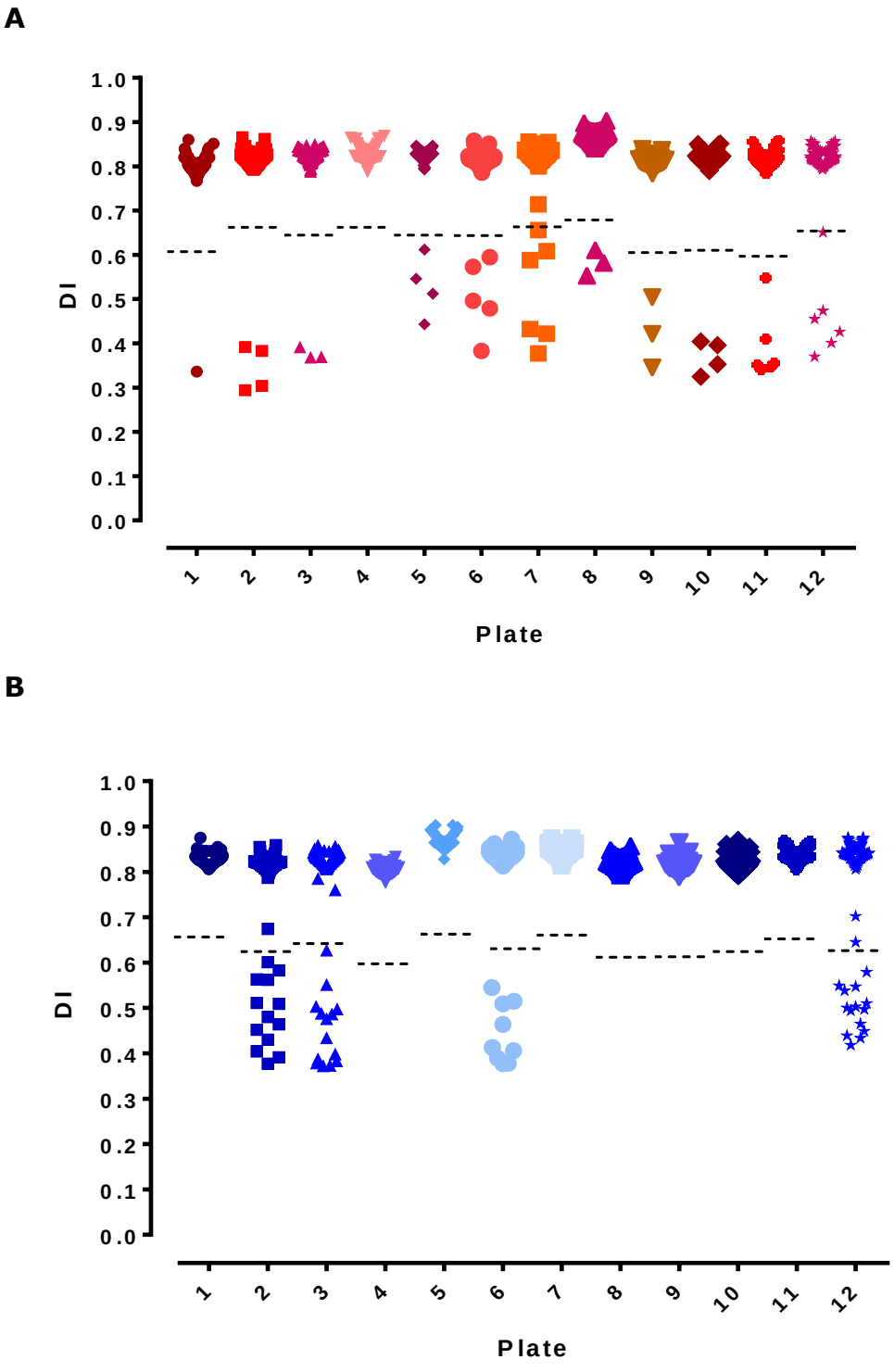


Figure 5.46: Screening results for *S. cerevisiae* MVD variants
(A) refers to Library A and (Bb) to Library B. DI values for the variants are indicated for each plate screened (1-12). The cut off limit set is indicated for each plate with a dashed black line.

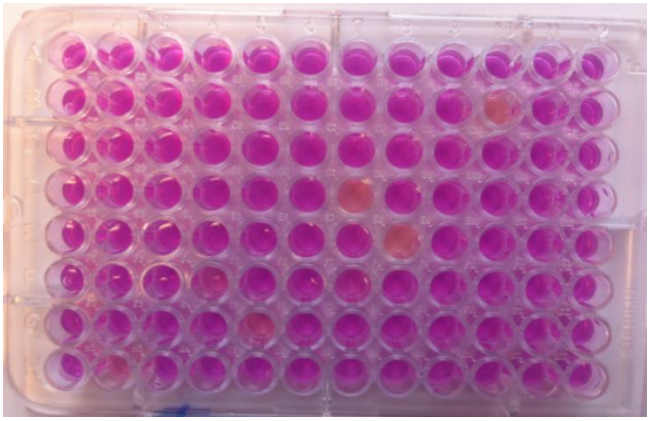


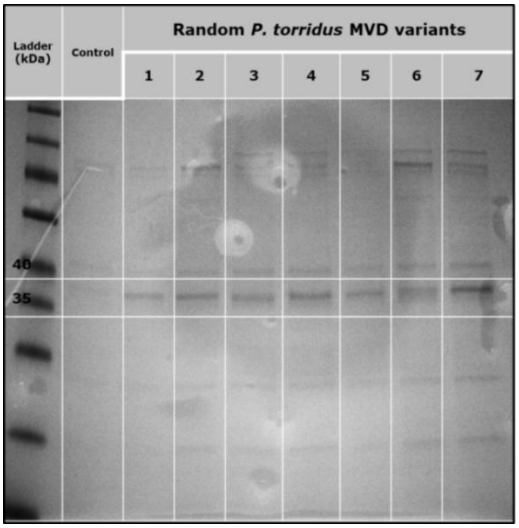
Figure 5.47: Example of a detection plate at the end of a Microresp^{TM} assay

To confirm that the release of CO_2 detected with Microresp^{TM} for the samples considered as positive hits was directly linked to MAA formation, the individual crude extracts were isolated from the plates and analysed by HPLC. Unfortunately, neither citramalic acid consumption nor MAA formation was observed for any of the samples. As further investigation, the 5 variants of each library that showed the best DI were tested at a larger scale (5 mL, 10 mM (*S*)-citramalic acid and 10 mM ATP). However, again, no MAA formation was observed, confirming the inability of the variants to catalyse the desired reaction.

The same five variants were also tested on 3-HIV (5 mM), in order to see if the mutagenesis affected enzyme activity towards this substrate. For *P. torridus*, activity was barely affected as all five variants showed a similar titre of isobutene formation (95-106% of wild type activity). For *S. cerevisiae* MVD, the variants showed a drastic decrease in activity instead (4-11% of wild type activity). This suggests that the residues selected in the design of *P. torridus* MVD libraries were not crucial for enzyme activity, as expected, whereas those selected for *S. cerevisiae* MVD played instead a more critical role. This was a risk to be aware of, since that three amino acids among the selected ones (Gly154, Ser155 and Ser208) were known to be believed to participate in the coordination of ATP (Barta *et al.*, 2012; Barta *et al.*, 2011; Byres *et al.*, 2007; Krepkiy and Miziorko, 2004; Krepkiy and Miziorko, 2005; Weerasinghe and

Dassanayake, 2009), but such a drastic effect was not expected. Mutagenesis almost completely killed enzymatic activity. It is possible that these residues have other functions, most likely to maintain enzyme stability. In fact, the analysis of random samples among the libraries showed solubility problems with *S. cerevisiae* MVD variants, and the enzyme was often not even observable on SDS-PAGE (Figure 5.48B, Lane 1-2-6-7). By contrast, enzyme variants of *P. torridus* MVD were always soluble and highly expressed (Figure 5.48A).

A



B

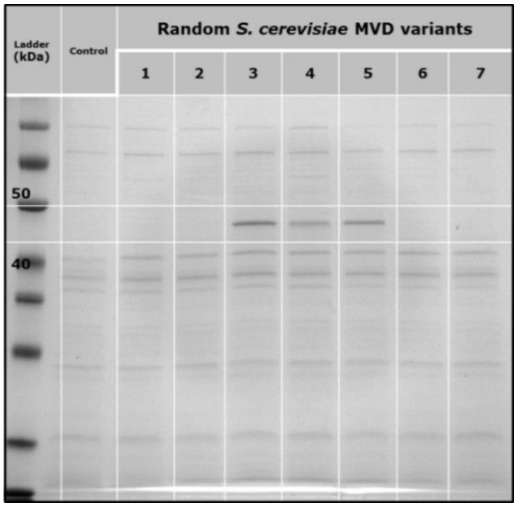


Figure 5.48: Analysis of expression of *S. cerevisiae* MVD and *P. torridus* MVD variants
SDS-PAGE analysis of seven random variants from Library A and B of *P. torridus* MVD (A) and *S. cerevisiae* MVD (B). (Control) refers to a sample containing crude extract obtained from cells not expressing the enzymes of interest. (Random MVD variants) refers to samples of crude extract containing MVD taken from random wells. The level of migration corresponding to *P. torridus* and *S. cerevisiae* MVD is marked between white lines.

5.5.2.3.3 Pool screening

As back up screening test for both *P. torridus* and *S. cerevisiae* MVD variants, the generated clones were also tested for direct production of MAA in pools *via* HPLC. Such a screening could not be performed for individual clones because of the impracticality of testing thousands of samples this way. *E. coli* BL21(DE3)pLysS was transformed with the plasmid libraries and then plated on several agar plates at a dilution that allowed growth of about 50-60 colonies per plate. For each library, 16 plates were prepared for a total number of 800-960 clones. For each plate, all the colonies were transferred together into the same fresh medium, induced and lysed. The crude extracts were then mixed with ATP and (*S*)-citramalic acid and incubated at 37°C. Since each pool consisted of 50-60 enzyme variants, a higher concentration (50 mM) of both ATP and citramalic acid was used in order to maximise the likelihood of finding activity even if given by a single variant of the pool. After 48 h, reaction mixtures were analysed by HPLC. Unfortunately, neither citramalic consumption nor MAA production was observed for any of the pools tested.

As for the single variants previously analysed, activity of the libraries on 3-HIV was almost unchanged for *P. torridus* MVD (97-102% variation) but was almost killed for *S. cerevisiae* MVD (2-8% of wild type activity).

5.5.3 Discussion

Attempts to modify *S. cerevisiae* and *P. torridus* MVD by protein engineering to enable enzymatic conversion of citramalic acid to MAA were performed. Since the enzymes were not able to bind citramalic acid (see section 5.4.2.3, Page 98), the strategy was to modify the binding pocket. For each enzyme, four residues were selected for mutagenesis and, according to the principles of ISM (Reetz and Carballeira, 2007; Reetz and Sanchis, 2008), they were split into two libraries of two residues each to minimise screening

efforts. Unfortunately, screening of the libraries generated did not lead to the discovery of any enzyme able to catalyse MAA formation and, therefore, the ISM principles were not applicable.

The screening assay used to monitor the activity of MVD, Microresp™, showed to be robust and reproducible, with an excellent sensitivity, and it also demonstrated to be applicable to different decarboxylases. This assay represents the first example of a general screening methodology for decarboxylases, and this could have a strong impact in the field of protein engineering. Nevertheless, when used for long incubation times (24-48 h), the Microresp™-based assay presents a limitation, a significant background noise that can mask or be confused for enzymatic activity. For this reason, cut off limits have to be set to minimise the risk of false positives but, even with this precaution, false positive were observed in the screening experiments.

An obvious explanation for the protein engineering strategy to be unsuccessful is hard to find. One reason could be that MVDs have too many residues around the binding pocket that are directly, or indirectly, involved in the catalytic mechanism. Mutagenesis in these regions could affect the function of some of these residues compromising the whole enzymatic mechanism. Aware of this, the residues mutated in the experiments were selected carefully in the attempt to avoid negative effects on enzyme activity. However, *S. cerevisiae* MVD activity was almost killed. By contrast, the activity of *P. torridus* MVD was not affected. This could be due to the fact that residues were selected based on a homology model and not on a real crystal structure. Therefore, it is possible that the selected residues were not exactly in the binding pocket as expected. Other rounds of mutagenesis on different residues could be performed, but the lack of time prevented this.

Another explanation for these enzymes not working on citramalic acid may reside in the difference between this substrate and the natural one, MVAPP. Citramalic acid presents a carboxylate moiety

which is absent on MVAPP, and this accounts for the inability of acting on this substrate. Since it was first believed that this was merely due to binding issues, it wasexpected to overcome the problem through modification of the binding pocket. However, it is also possible that this difference affect the reaction mechanism in a different way. Provided that the enzymes are able to bind citramalic in the active site, it may be that the carboxylate of the substrate interferes with the catalytic residues, impeding the reaction. The carboxylate is very close to the hydroxyl group that requires deprotonation before the substrate can be phosphorylated and decarboxylated (Figure 5.49A). The carboxylate may interact with the catalytic residues involved in such a deprotonation (usually Asp), stopping the reaction at the earliest stage. If this is true, it is highly unlikely that any protein engineering strategy could be successful. Testing an alkyl ester of citramalic acid, where the carboxylate is esterified with an alkyl chain, would be an interesting experiment to perform, since the effect of the carboxylate should be neutralised (Figure 5.49B). However, citramalic acid esters are very challenging to synthesise either chemically or biologically, because it is difficult to selectively esterify one carboxylate over the other one. For this reason, it was not able to proceed with this investigation.

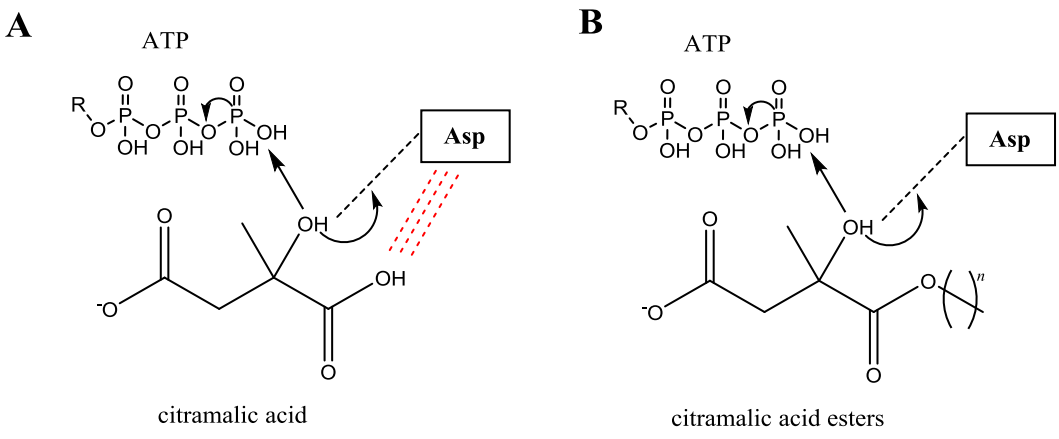


Figure 5.49: Possible interference of citramalic acid with the catalytic residues of MVDs

Unfortunately, with these experiments, it was demonstrated that protein engineering on MVD to develop an enzyme able to catalyse the conversion of citramalic acid to MAA is not very promising. Hence, other strategies and solution must be explored.

On a separate note, the multiple sequence alignment of MVD homologues raised some interesting considerations. The sequence identity and similarity level of *P. torridus* MVD with other more characterised MVDs is very low (18% and 36% with *S. cerevisiae*). The most significant differences are observed for regions suggested to be involved in the binding of the diphosphate moiety of MVAPP (Figure 5.39 - position 17-21, 192-196). This may explain why *P. torridus* MVD was reported to have a wider substrate specificity than other MVDs towards substrates that do not present that diphosphate moiety (Marliere, 2010; Marliere *et al.*, 2011). Moreover, another significant difference can be viewed at position 283 (Figure 5.39). Whereas all the other MVDs carry an Asp, *P. torridus* MVD presents a Thr. Since Asp is known to be crucial for catalysis by functioning as the deprotonating acid to allow phosphorylation of the substrate (Figure 5.20, Page 92), this result is very curious because it may suggest a slightly different deprotonating mechanism not described before for this class of enzymes. Another explanation could be that this catalyst is different to classical MVDs. Although *P. torridus* MVD is annotated as an MVD and it has been referred as MVD in different patents (Marliere, 2011; Marliere, 2010; Marliere *et al.*, 2011), a full characterization has never been reported. These observations led me to perform further investigation on this enzyme to explain the differences and to gain more insights into its potential use for MAA formation (see next chapter).

5.6 Investigation on the real biological function of *P. torridus* MVD

5.6.1 Introduction

Since *P. torridus* MVD is still poorly known compared to other MVDs, it was decided to purify it and to perform a detailed characterisation. After sequence analysis, the enzyme was tested on the putative natural substrate MVAPP and on other potential substrates. The enzyme's ability to produce isobutene from 3-HIV and its potential to produce other valuable small alkenes were also investigated.

5.6.2 Results

5.6.2.1 Sequence and structure analysis

Initially, a bioinformatics approach was used to gain more insights into *P. torridus* MVD. A BLASTp search for homologues and a phylogenetic analysis revealed that proteins with the highest sequence identity and similarity are found mostly in archaea, with higher values with enzymes from *Ferroplasma acidarmanus*, *Thermoplasma volcanium* and *Thermoplasma acidophilum*, all members of the Thermoplasmatales order (Figure 5.50; Table 9.5, Page 228). As already observed, *P. torridus* MVD shares a sequence identity and similarity of only 18% and 36% with *S. cerevisiae* MVD but also of only 21% and 41% with *S. epidermidis* MVD (Barta *et al.*, 2011). This significant difference opens the possibility that *P. torridus* MVD, and the other Thermoplasmatales homologues, may be unrelated to classical MVDs. To further investigate this hypothesis, multiple sequence alignment was made between Thermoplasmatales homologues and classical MVDs to compare important conserved regions and residues known to be involved in substrate binding and enzyme activity (Figure 5.51).

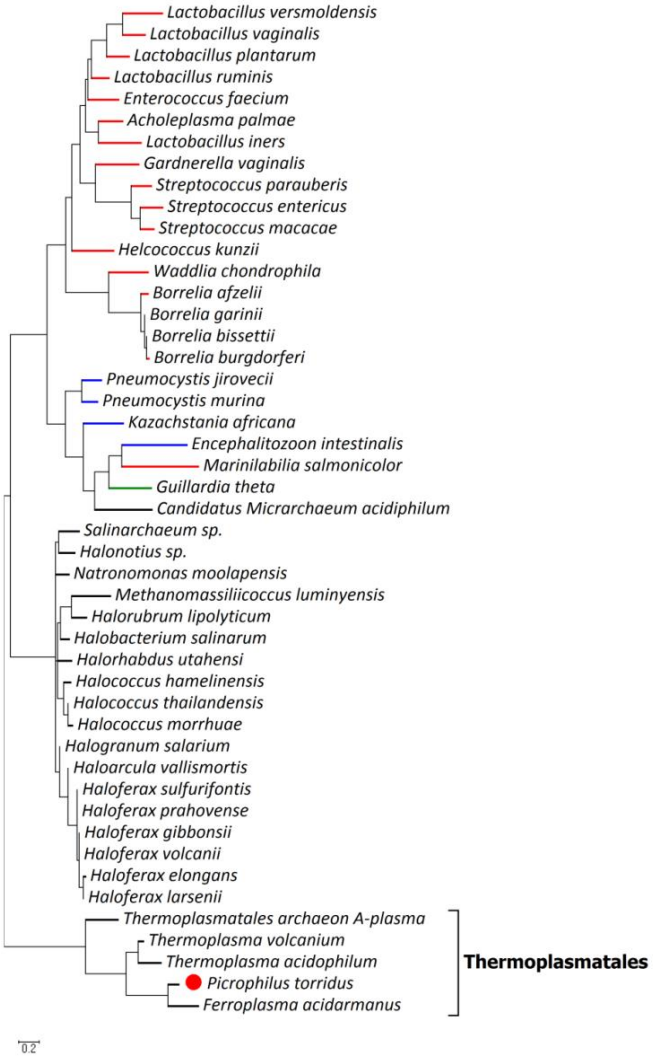


Figure 5.50: Phylogenetic analysis
P. torridus MVD is highlighted with a red circle. Organisms highlighted in red belong to the Bacteria kingdom, in blue to the Fungi kingdom, in green to the Chromalveolata kingdom, in black to the Archaea kingdom. The black bracket indicates archaea belonging to the Thermoplasmatales order.

The alignment comprised the same enzymes used to analyse residues for mutagenesis (see section 5.5.2.2.1, Page 127) but with the addition of the other Thermoplasmatales homologues. As for *P. torridus* MVD, also in all the other Thermoplasmatales homologues the regions for ATP binding and interaction with the C1 carboxylate of MVAPP are similar to classical MVDs (Figure 5.51, Position 12, 101-11, 144) but strong differences occur in the regions involved in the binding of the diphosphate moiety of MVAPP or directly involved in the catalytic mechanism (Figure 5.51, Position 17-21, 72, 139-141, 192-196, 283). This information suggests that

Thermoplasmatales carry a unique "MVD-like" protein which has significant differences in the amino acids implicated in binding the diphosphate moiety of MVAPP and that may be unrelated. Therefore, it was decided to investigate *P. torridus* MVD substrate specificity and reaction selectivity.

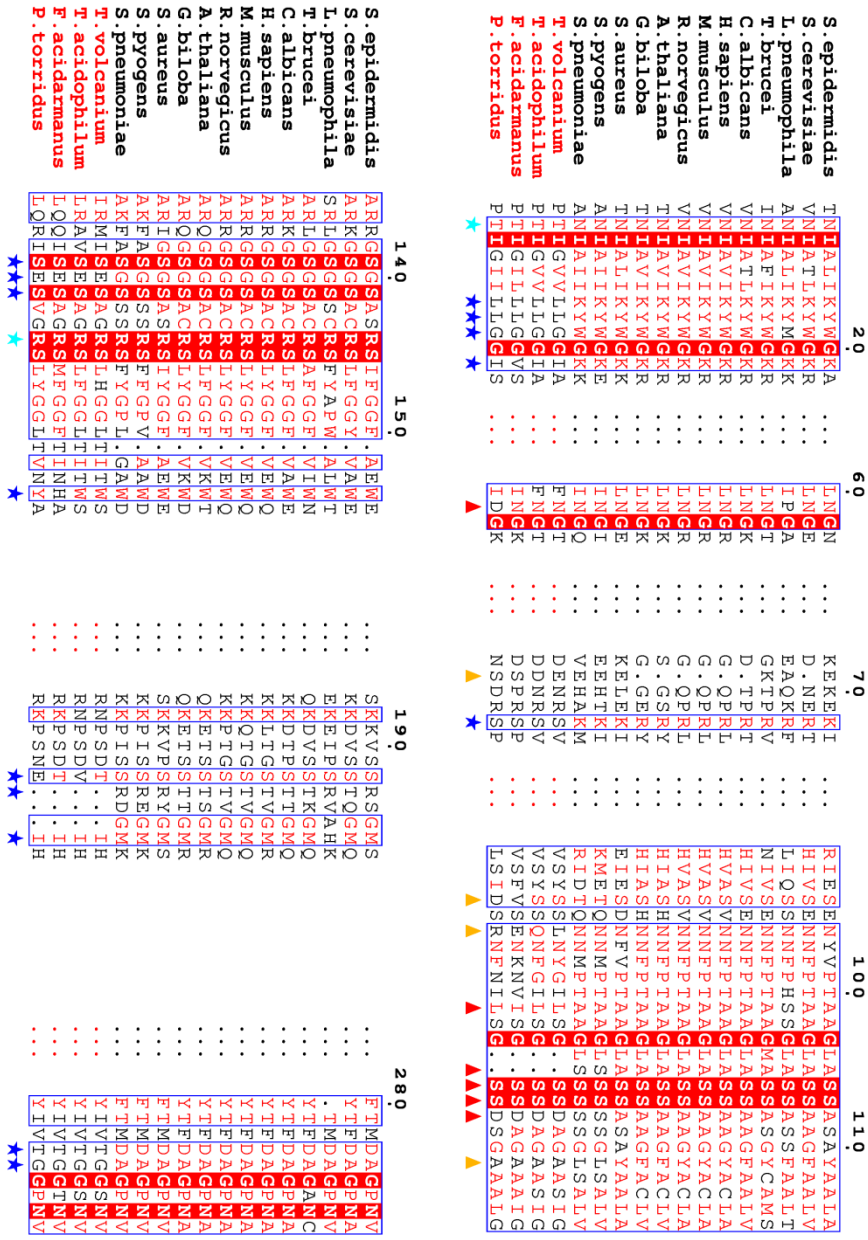


Figure 5.51: Multiple sequence alignment of MVDs and Thermoplasmatales homologues
Conventional MVDs are indicated in black and "MVD-like" proteins from the Thermoplasmatales order in red. Only highly conserved regions (framed in blue) are depicted. Conservation of residues is denoted by colours: white amino acids highlighted in red are strictly conserved; red amino acids not highlighted are high in similarity. The numbering annotation is based on *S. epidermidis* MVD. Stars denote residues involved in MVAPP interaction (cyan for C1 carboxylate group, blue for diphosphate moiety and hydroxyl group). Triangles denote residues involved in ATP interaction (red for phosphate groups, orange for purine ring and ribose).

5.6.2.2 The search for the natural substrate

Although *P. torridus* MVD had been previously referred as an MVD (Marliere, 2011; Marliere, 2010; Marliere *et al.*, 2011), its direct activity on MVAPP was never reported. Even in the experiments to verify that the enzyme was expressed in active form, 3-HIV was used as a substrate and not MVAPP. Therefore, the first step was to verify its activity on MVAPP. *P. torridus* MVD was expressed with an N-terminal his-tag (6xHis) and purified on a Ni Sepharose column. The protein was at least 95% pure (Figure 5.52).

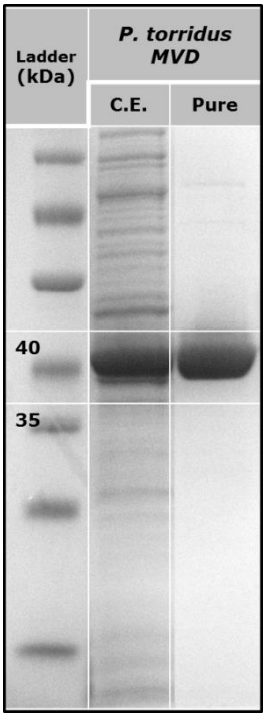


Figure 5.52: Purification of *P torridus* MVD.
SDS-PAGE analysis of *P. torridus* MVD in crude extract (C.E.), and after purification on a Ni Sepharose column (Pure). The level of migration corresponding to MVD is marked between white lines.

Activity was investigated using a coupled enzymatic assay already in use for the analysis of MVDs (Alvear *et al.*, 1982; Barta *et al.*, 2012; Barta *et al.*, 2011; Krepkiy and Miziorko, 2005; Krepkiy and Miziorko, 2004). Surprisingly, no activity was detected when MVAPP was tested as the substrate. Therefore, two other substrates were also tested: (*R*,*S*)-mevalonate-5-phosphate (MVA-5P) and (*R*)-mevalonate (MVA). MVA-5P is the substrate for MVA-5P decarboxylases (PMD), which have recently been discovered in

Haloferax volcanii (Vannice *et al.*, 2014) and *Roseiflexus castenholzii* (Dellas *et al.*, 2013; Vannice *et al.*, 2014). Unlike MVAPP and MVA-5P, MVA lacks any phosphate moiety. Whereas there was no detectable activity on MVA-5P, activity with MVA was detected at a rate of 3.01 Units·mg^{-1} (Table 5.15). As a comparison, *S. cerevisiae* MVD was also expressed and purified in the same way as *P. torridus* MVD, and tested on the same substrates. This enzyme was active only on MVAPP and not on MVA-5P or MVA (Table 5.15). This result suggested that *P. torridus* MVD was probably not an MVD but a MVA decarboxylase instead.

Table 5.15: Kinetic constants of *S. cerevisiae* MVD, *P. torridus* MVD, and PMD from *H. volcanii* and *R. castenholzii* [a]

Enzyme	Activity (Units · mg^{-1})	K_m [b] (μM)	$K_{m\ ATP}$ (μM)	K_{cat} (s^{-1})	K_{cat}/K_m [b] ($s^{-1}\cdot\mu M^{-1}$)
P. torridus MVD	3 ± 0.1	131 ± 18 (MVA)	23 ± 0.1	1.9	1.5×10^{-2} (MVA)
S. cerevisiae MVD(this paper)	7 ± 0.7	133 ± 28 (MVAPP)	61 ± 6	5.4	4×10^{-2} (MVAPP)
S. cerevisiae MVD(Krepkiy and Miziorko, 2004; Krepkiy and Miziorko, 2005)	6.4 ± 0.2	123 ± 22 (MVAPP)	61 ± 6	4.9	4×10^{-2} (MVAPP)
H. volcanii PMD(Vannice *et al.*, 2014)	5.6 ± 0.1	159 ± 15 (MVA-5P)	289 ± 15	3.5	2.2×10^{-2} (MVA-5P)
R. castenholzii PMD(Dellas *et al.*, 2013)	*n.a.*	152 ± 38 (MVA-5P)	*n.a.*	1.7	1.1×10^{-2} (MVA-5P)

[a] Activity values refer to the rate of ADP formation coupled with of NADH oxidation .Values for *P, torridus* MVD and *S. cerevisiae* MVD were calculated from enzyme assays at 30°C in 50 mM Tris-HCl buffer, pH 7.5, with the addition of 10 mM $MgCl_2$, 20 mM KCl and 10 μg of pure enzyme. Assays were done in triplicate and the means and errors for the kinetic constants are shown. Previous data for *S. cerevisiae* MVD, for *H. volcanii* PMD, *R. castenholzii* PMD are taken from the literature. *n.a* not available

[b] The values are given for the respective substrate, in parentheses, for each of the enzyme (MVAPP, MVA-5P, or MVA)

In order to understand whether MVA was the natural substrate of *P. torridus* MVD, the kinetic parameters for MVA phosphorylation with MVAPP phosphorylation by *S. cerevisiae* MVD were measured and compared (Table 5.15). The kinetic parameters for *S. cerevisiae* MVD were very similar to those reported previously (Krepkiy and Miziorko, 2005; Krepkiy and Miziorko, 2004). The magnitude of the kinetic constants and activities of *P. torridus* MVD with MVA were similar to those of *S. cerevisiae* MVD with MVAPP, consistent with the hypothesis that MVA is the natural substrate of *P. torridus* MVD, as MVAPP is for *S. cerevisiae* MVD. In addition, the kinetic

parameters reported in the literature for the newly discovered PMDs (Dellas *et al.*, 2013; Vannice *et al.*, 2014) are also similar in magnitude (Table 5.15).

5.6.2.3 The new enzyme classification

To be sure that *P. torridus* MVD was a MVA decarboxylase, further experiments were needed, since the coupled enzyme assay only monitors ADP release and not the actual decarboxylation of the intermediate. The expected reaction product would be 3-methylbut-3-en-1-ol (isoprenol) if MVA is phosphorylated and then decarboxylated. Therefore, attempts were made to detect isoprenol by Gas Chromatography – Mass Spectrometry (GC-MS). Purified *P. torridus* MVD was mixed with a high concentration of MVA (5 mM) and the presence of isoprenol was checked at intervals (24, 48 and 96 h). Even after a long incubation time (96 h), isoprenol was not observed, suggesting that *P. torridus* MVD does not catalyse MVA decarboxylation *per se*.

Therefore, *P. torridus* MVD was tested for the formation of a phosphorylated product from MVA, analysing the reaction product by Electrospray Ionization - Mass Spectrometry (ESI-MS). An ion with m/z-227 was detected only in the samples where *P. torridus* MVD was added to the reaction mixture, which corresponded to the mass of the expected phosphorylated MVA (Figure 5.53B). In the control without added enzyme, this ion was not present. Instead, a 147 m/z ion was among those with the highest intensity, corresponding to unreacted MVA (Figure 5.53A). A higher background noise was observed in the MS trace for the control compared to the test. This was probably due to a lower response of MVA in the ESI-MS analysis compared to the phosphorylated MVA. Unreacted MVA was observed also in samples where the enzyme was added, meaning that full conversion was not reached (Figure 5.53B). As a confirmation that the 227 m/z ion corresponded to phosphorylated MVA, MS-MS analysis of the isolated product

generated peaks corresponding to phosphate ions ($H_2PO_4^-$, 97 and PO_3^-, 79 m/z) (Figure 5.53C).

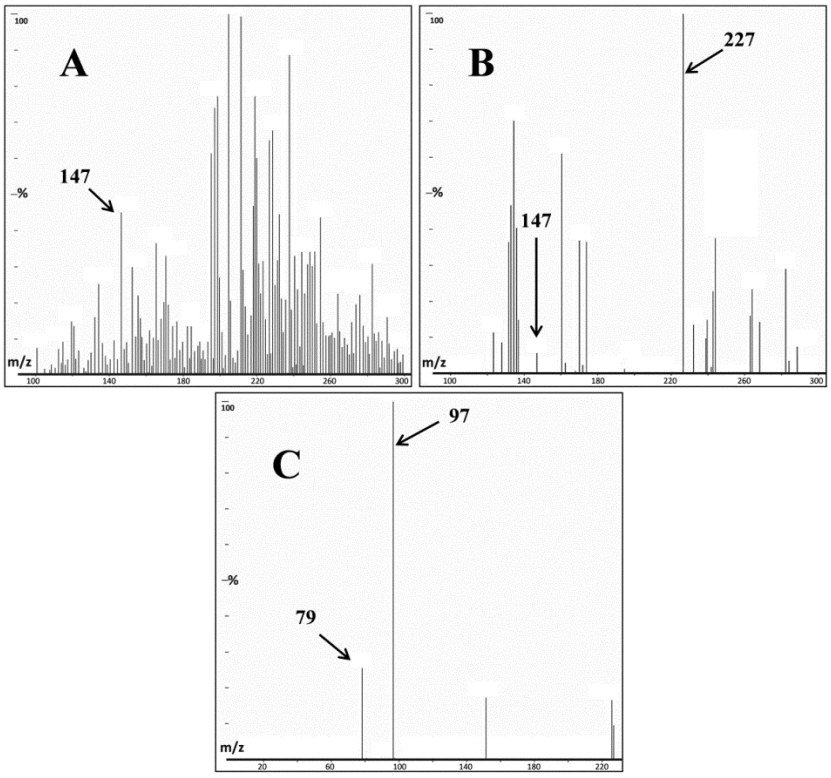
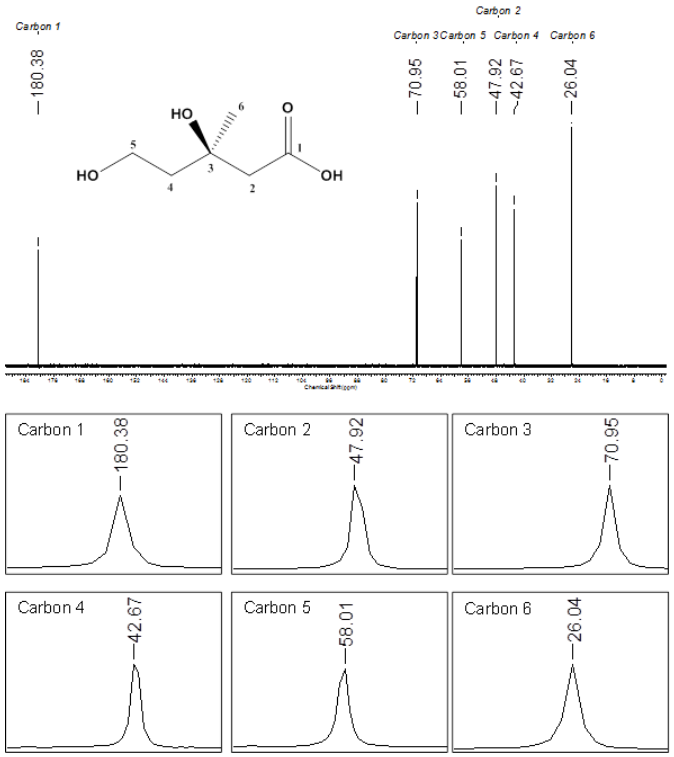


Figure 5.53: ESI-MS analysis of *P. torridus* MVD product
All spectra were collected in negative ion mode. The mass spectra are shown. (A) Control sample containing MVA but without added enzyme; (B) Sample containing MVA and *P. torridus* MVD; (C) MS-MS spectrum of 227 m/z ion yielding phosphate ions (97 and 79 m/z). The reaction was performed in 50 mM Tris-HCl buffer, pH 7.5, with the addition of 10 mM $MgCl_2$ and 20 mM KCl. Samples were analysed after 24 h.

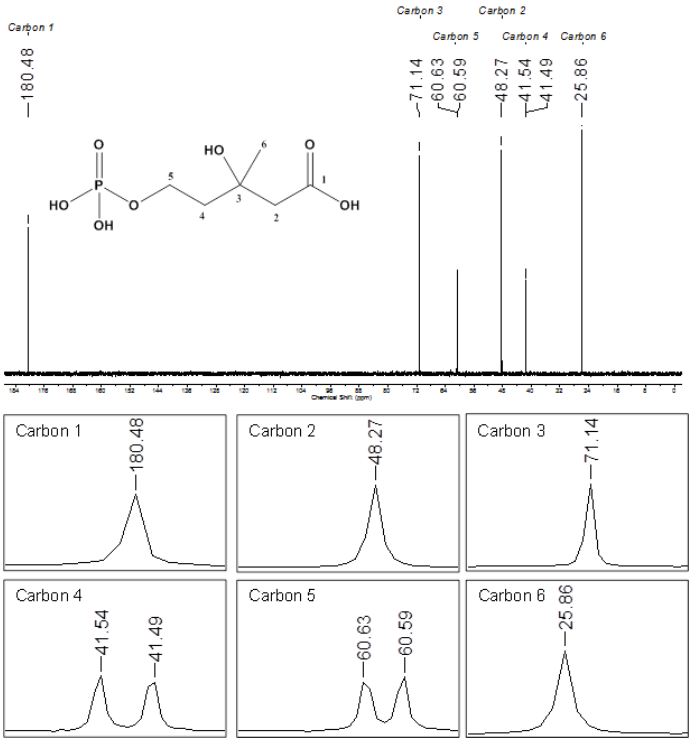
To discriminate between MVA-5P and MVA-3P, both possible phosphorylation products from MVA, ^{13}C NMR was used and the coupling between ^{13}C and ^{31}P was analysed. Any coupling would have resulted in fine structure (doublets) associated with the signals for each carbon atom within three bonds from ^{31}P. The ^{13}C NMR spectrum of authentic MVA showed six singlet peaks whereas the spectrum of authentic MVA-5P showed splitting at carbons 4 and 5 as expected (Figure 5.54A and B). By contrast, ^{13}C NMR spectrum of the product mixture after phosphorylation of MVA by pure *P. torridus* MVD showed splitting at carbons 2-4 and 6, consistent with phosphorylation at position 3 (Figure 5.54C). Thus, analysis of the sequence, substrate range and reaction products demonstrated that *P. torridus* MVD was annotated incorrectly, and that was actually a

mevalonate-3-kinase (M3K) that acts on MVA to produce MVA-3P (Figure 5.55A). Therefore, the enzyme should be renamed as *P. torridus* M3K. From now on, it will be described as such throughout the text.

A



B



C

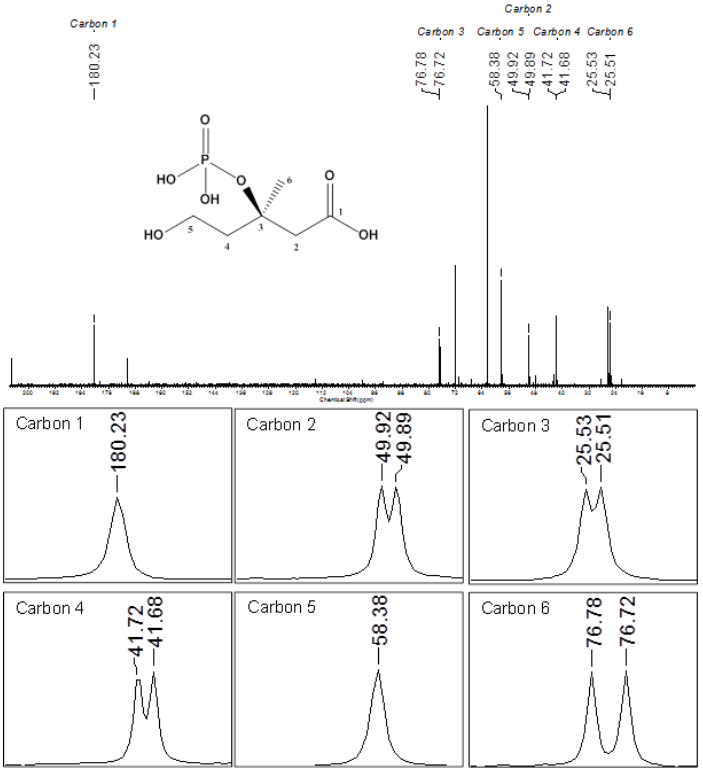


Figure 5.54: Full NMR spectra.
^{13}C NMR spectra of (*R*)-mevalonate (A), (*R*,*S*)-mevalonate-5-phosphate (B) and *P. torridus* MVD product (C). The structure of the compound is shown at the top, and the peaks corresponding to carbons 1to 6 are shown at the bottom. Values are in parts per million (ppm). Concentration of (*R*)-mevalonate (MVA) and (*R*,*S*)-mevalonate-5-phosphate (MVA-5P) was 40 mM. For *P. torridus* MVD product, the reaction was performed as previously described (Vinokur *et al.*, 2014a) using MVA as a substrate at a concentration of 40 mM, and samples were analysed after 24 h. Spectra were acquired on a 500 MHz Bruker AV(III)500 spectrometer equipped with a cryoprobe. Data were processed and analysed using ACD/NMR Processor version 12.01.

5.6.2.4 Formation of isobutene

P. torridus M3K was identified as a preferred enzyme for the production of isobutene from 3-HIV (Marliere, 2010; Marliere *et al.*, 2011), but it was believed to act as an MVD and not as an M3K. More recently, higher yields for production of isobutene were obtained when *P. torridus* M3K was mixed with other MVDs (Delcourt *et al.*, 2013; Marliere *et al.*, 2012). To explain this, it was proposed that *P. torridus* M3K (still considered as an MVD) catalyses phosphorylation of 3-HIV faster than the subsequent decarboxylation. It was further proposed that the phosphorylated intermediate could be released and that the additional mixture of

other MVDs could catalyse decarboxylation of this intermediate to isobutene. However, the authors did not further investigate the hypothesis, and kept considering the enzyme as an MVD. From the results, it is clear now that the enzyme is an M3K, and not a true decarboxylase. Therefore, the production of isobutene depends only on the efficient formation of an unstable intermediate, 3-phosphoisovalerate (3-PIV; Figure 5.55B). In the absence of other enzymes, this intermediate decarboxylates non-enzymatically to isobutene. The presence of other MVDs may provide an overall increase in efficiency.

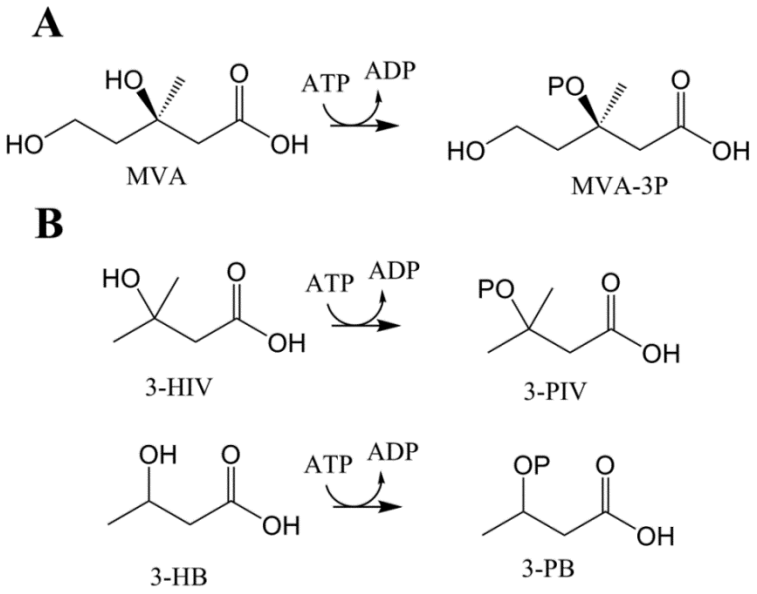


Figure 5.55: M3K reaction scheme.
(A) The natural substrate of M3K is (*R*)-mevalonate (MVA), that is converted to (*R*)-mevalonate-3-phosphate (MVA-3P). (B) *P. torridus* M3K is also active on 3-hydroxyisovalerate (3-HIV) and 3-hydroxybutyrate (3-HB) to produce 3-phosphoisovalerate (3-PIV) and 3-phosphobutyrate (3-PB), respectively

Since a full description of this reaction was still not available, it was decided to investigate it in more detail. Activity of *P. torridus* M3K on 3-HIV was first evaluated. Compared to the natural substrate, MVA, the catalytic efficiency was very low, with a 20-fold higher K_m and a 1000-fold lower K_{cat}/K_m (Table 5.16). Nevertheless, the K_m and K_{cat}/K_m were lower (2.35 mM) and higher (4.5×10^{-5} s^{-1} μM^{-1}), respectively, than those disclosed previously (9.17 mM and 2×10^{-5} s^{-1} μM^{-1}) (Marliere *et al.*, 2012). Next, isobutene formation

was compared using whole cells expressing *P. torridus* M3K, crude extract containing *P. torridus* M3K, and pure *P. torridus* M3K (Table 5.17). Production rates were compared with *S. cerevisiae* MVD under the same conditions, since it also catalyses isobutene formation from 3-HIV (Gogerty and Bobik, 2010).

Table 5.16: Kinetic constants of *P. torridus* M3K on MVA, 3-HIV, 3-HB and 3-HP [a]

Substrate	**Activity** (Units·mg^{-1})	**K_m** (μM)	**K_{cat}** (s^{-1})	**K_{cat}/K_m** (s^{-1}·μM^{-1})
Mevalonate (MVA)	3 ± 0.1	131 ± 18	1.9	1.5×10^{-2}
3-hydroxyisovalerate (3-HIV)	0.17 ± 0.01	2348 ± 574	0.1	4.25×10^{-5}
3-hydroxybutyrate (3-HB)	0.14 ± 0.01	3163 ± 637	0.1	2.9×10^{-5}
3-hydroxypropionate (3-HP)	*n.d.*	*n.d.*	*n.d.*	*n.d.*

[a] Activity values refer to the rate of ADP formation coupled with of NADH oxidation. Values are calculated at 30°C in 50 mM Tris-HCl buffer, pH 7.5, with the addition of 10 mM $MgCl_2$, 20 mM KCl and 50 μg of purified *P. torridus* M3K (for MVA, purified *P. torridus* M3K added was 10 μg). Assays were done in triplicate and the means and errors for the kinetic constants are shown. *n.d.* not detected.

As expected (Gogerty and Bobik, 2010; Marliere, 2010; Marliere *et al.*, 2011), both *P. torridus* M3K and *S. cerevisiae* MVD catalysed formation of isobutene (Table 5.17). In controls without ATP, isobutene formation was similar to those observed for controls without enzyme. This confirmed that the reactions catalysed by the enzymes were ATP-dependent. *P. torridus* M3K catalysed isobutene formation at rates consistently higher than *S. cerevisiae* MVD under all assay conditions. With the pure enzyme, an 11-fold improvement was observed (Table 5.17). Very interestingly, a further 110-fold enhancement (from 26 to 2880 pmol min^{-1} $mg_{(protein)}{}^{-}1$) was observed for purified *P. torridus* M3K when the temperature of the assay was increased from 30 to 50°C (Table 5.17). It is possible that this was due to an increased rate of spontaneous decarboxylation of 3-PIV. It is worth noting that the rates of isobutene formation obtained using *P. torridus* M3K were even higher than reported previously using enhanced variants of

S. cerevisiae MVD obtained with protein engineering tools (Gogerty and Bobik, 2010).

Table 5.17: Isobutene formation rates for *P. torridus* M3K and *S. cerevisiae* MVD [a]

Enzyme	Whole cells ($pmol \cdot min^{-1} \cdot g_{(cells)}^{-1}$)	Crude extract ($pmol \cdot min^{-1} \cdot mg_{(total\ protein)}^{-1}$)	Purified protein ($pmol \cdot min^{-1} \cdot mg_{(protein)}^{-1}$)	
			30°C	50°C
P. torridus M3K	507 ± 137	261 ± 19	26 ± 2	2880 ± 140
S. cerevisiae MVD	7.4 ± 1.5	1.6. ± 0.3	2.3 ± 1.2	*n.a.*

[a] The concentration of 3-HIV used was 50 mM. For the crude extract and purified enzyme experiments, 40 mM ATP was added. Assays were done in triplicate and the means and errors for formation rates are shown. Isobutene was quantified by GC-MS analysis of the headspace. Due to the very low aqueous solubility of isobutene (Yaws, 2003), its concentration in the liquid phase was considered negligible and was not taken into consideration.

5.6.2.5 Formation of other gaseous alkenes

Finally, in order to test the use of *P. torridus* M3K for production of other small, industrially useful alkenes, its activity was investigated on two other substrates, 3-hydroxypropionate (3-HP) and 3-hydroxybutyrate (3-HB). These substrates have the potential to be converted to ethene and propene (Obenaus *et al.*, 2000) if they can be phosphorylated at the 3-position, and if the intermediates are sufficiently unstable to decarboxylate spontaneously. Although the enzyme was able to phosphorylate 3-HB (Figure 5.55B) with kinetics similar to those for 3-HIV (Table 5.16), no propene was detected by GC-MS using either crude extract or purified enzyme, even at a temperature of 50°C. This suggests that, as for MVA, the phosphorylated intermediate is too stable to decarboxylate. As for 3-HP, the enzyme was not even able to catalyse its phosphorylation (Table 5.16). This suggests that one or two methyl substituents are required at position C3 to enable binding of the substrate to the enzyme and/or for catalytic activity.

5.6.3 Discussion

It was demonstrated that the *P. torridus* enzyme (GenBank (Benson *et al.*, 2013) accession number AAT43941), so far classified as mevalonate diphosphate decarboxylase, is actually a

mevalonate-3-kinase, since it acted on MVA to produce MVA-3P (Figure 5.55A). This intermediate appeared to be stable, since no decarboxylation of MVA was observed and MVA-3P could be detected in the reaction mixture. Nevertheless, the enzyme can still function as a "decarboxylase-like" catalyst when the phosphate intermediate generated is unstable. In fact, it was able to catalyse isobutene formation from 3-HIV through formation of the unstable 3-PIV (Figure 5.55B).

Isobutene is a small, highly reactive molecule, used extensively as a platform chemical to manufacture a wide variety of products (Obenaus *et al.*, 2000; vanLeeuwen *et al.*, 2012) including fuel additives, rubbers and speciality chemicals. The global demand for isobutene produced from petrochemical sources was estimated to be around 10 million metric tonnes per year with a market value higher than 15 billion euros (de Guzman, 2011; OECD, 2003). Sustainable conversion of renewable feedstocks to isobutene using engineered pathways is, therefore, an important target. Global Bioenergies (Evry, France), company which is using MVDs and "MVD-like" proteins for isobutene production, is currently performing pilot plant studies (Global Bioenergies, 2014). The discovery that the *P. torridus* homologue is in reality an M3K could have a beneficial effect in the field. This enzyme showed a 5-fold increased production rate (507 pmol min^{-1} $g_{(cells)}^{-1}$) compared to the rate previously reported with *S. cerevisiae* MVD (98 pmol min^{-1} $g_{(cells)}^{-1}$ with whole cells) (Gogerty and Bobik, 2010). Moreover, the discovery of a new class of enzyme opens the possibility to screen other, related enzymes for isobutene formation or to undertake protein engineering to improve the production rates still further.

As for the potential of *P. torridus* M3K to be used for the production of other gaseous alkenes such as ethene and propene, unfortunately, no product formation was observed. The enzyme was able to phosphorylate 3-HB (Figure 5.55B; Table 5.16) but, under the tested conditions, the 3P intermediate did not spontaneously decarboxylate to propene. This could be due to the

3-phosphobutyrate (3-PB) intermediate being more stable than 3-PIV. The absence of an electron-donating methyl group on 3-PB could reduce the electron density relative to the tertiary carbon of the intermediate and slow down the rate of the decarboxylation, thought to occur via a carbocation intermediate (Dhe-Paganon *et al.*, 1994).

In regard to the structural reasons why *P. torridus* M3K is not able to catalyse the decarboxylation of MVA or, as classical MVDs, of MVAPP, a conclusive explanation is still not possible. In classical MVDs, the phosphorylation of MVAPP results in formation of a transient tertiary phosphorylated intermediate that quickly releases inorganic phosphate with concomitant decarboxylation (Barta *et al.*, 2011; Byres *et al.*, 2007; Voynova *et al.*, 2008) or, even, in a concerted manner (Lefurgy *et al.*, 2010). Thus, the difference in substrate specificity (MVA *vs*. MVAPP) may account for the differences in the ability of *P. torridus* M3K and MVDs to act as decarboxylases. Another explanation is that this difference is due to the diversity in amino acid composition in regions suggested to be involved in the catalytic mechanism (Figure 5.51; position 281-284). In this region, classical MVDs have a strictly conserved Asp, that is needed to deprotonate the hydroxyl group, allowing subsequent phosphorylation (Barta *et al.*, 2012). In *P. torridus* M3K, Asp is substituted by Thr. This may suggest a slightly different deprotonating mechanism, perhaps involving the Thr or surrounding residues.

It is important to note that these differences, including those in regions that in MVDs are involved in the binding of MVAPP (Figure 5.51; position 17-21, 72, 139-141, 192-196), are constant for all the Thermoplasmatales homologues used in the analysis. Therefore, although still annotated as MVD, they could all be M3Ks. While these experiments were being performed, similar studies on the homologue from *T. acidophilum* were disclosed (Azami *et al.*, 2014; Vinokur *et al.*, 2014a; Vinokur *et al.*, 2014b), confirming that this enzyme was also an M3K and not an MVD. These studies also

confirmed that that the amino acid differences described above could be critical for the ability of the enzyme to catalyse phosphorylation rather than decarboxylation (Vinokur *et al.*, 2014b). Moreover, it was also demonstrated that M3K is part of a completely new type of MVA pathway (Azami *et al.*, 2014; Vinokur *et al.*, 2014a).

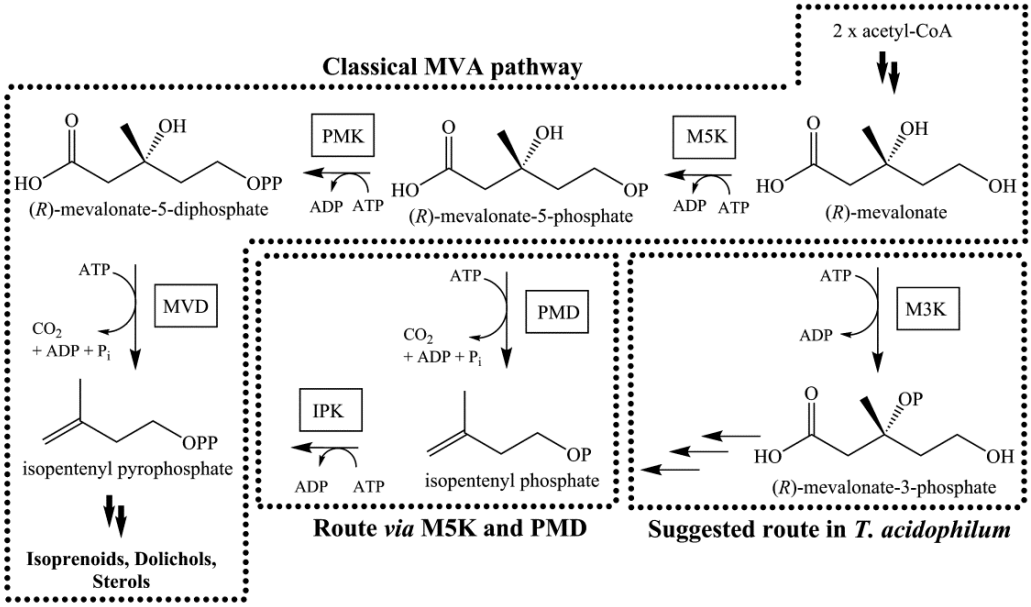


Figure 7

Figure 5.56 Classical MVA pathway and modified routes.
The classical MVA pathway and the modified routes (*via* M5K and PMD and the newly suggested route in *T. acidophilum*) are specifically indicated. Enzyme abbreviations are as follows: M5K, mevalonate-5-kinase; PMK, phosphomevalonate (mevalonate-5-phosphate) kinase; MVD, mevalonate-5-diphosphate decarboxylase; PMD, phosphomevalonate (mevalonate-5-phosphate) decarboxylase; IPK, isopentenyl phosphate kinase; M3K, mevalonate-3-kinase.

In the classical MVA pathway, MVA is double phosphorylated to MVAPP and subsequently converted to isopentenyl pyrophosphate (IPP) by MVD (Figure 5.56). In the newly described alternative, M3K catalyses the initial step to produce MVA-3P from MVA (Figure 5.56), and MVA-3P is subsequently converted to (*R*)-mevalonate-3,5-bisphosphate (MVA-3,5PP) by another new enzyme, MVA-3P-5-kinase (M3P5K; GenBank (Benson *et al.*, 2013) accession number CAC11895) (Azami *et al.*, 2014; Vinokur *et al.*, 2014a). The enzymes catalysing conversion of MVA-3,5PP to isopentenyl phosphate (IP) have not yet been identified, and this provides an

excellent target for future research. IP is finally converted to IPP via IP kinase (IPK) (Chen and Poulter, 2010) (Figure 5.56). It was now shown that an M3K is also present in *P. torridus*. Since a strongly similar counterpart of M3P5K is also found in *P. torridus* (Table 9.6 – Page 229), *P. torridus* is also likely to contain this new type of MVA pathway (Figure 5.56).

It should be noted that this pathway is distinct from the mevalonate-5-kinase (M5K) pathway found in *Haloferax volcanii* (Vannice *et al.*, 2014) and *Roseiflexus castenholzii* (Dellas *et al.*, 2013) (Figure 5.56). Neither of the key enzymes, (M5K and PMD) are present in *T. acidophilum* (Chen and Poulter, 2010) and there is no experimental or sequence homology evidence to support their presence in *P. torridus*. Therefore, *P. torridus* is proposed to produce IPP *via* the newly discovered M3K pathway.

From a microbiological/evolutionary point of view, this study raises important questions on the presence of alternative MVA pathways in archaea. It is even possible that variant MVA pathways are present amongst members of all kingdoms in nature (Dellas *et al.*, 2013). Many plant, bacterial and fungal enzymes are believed to participate in the classical MVA pathway, but, like *P. torridus* M3K, are still only computationally annotated and need to be investigated experimentally. Discovery of new enzymes is an important target, since this may yield biocatalysts with improved properties. Therefore, discovery of new enzymes of the MVA pathways could aid in improved metabolic engineering for efficient, sustainable bioproduction of chemicals and fuels.

5.7 Summary of results

The presented work investigated the use of decarboxylases for MAA production through conversion of itaconic, mesaconic, citraconic or citramalic acid.

- Eight candidate enzymes, belonging to different subclasses of decarboxylases (Table 5.18), were accurately selected based on binding and mechanistic requirements

Table 5.18: Selected enzyme for MAA formation

Enzyme
Phenylacrylic acid decarboxylase (PAD1)
Phenylacrylic acid decarboxylase (PadA1)
Ferulic acid decarboxylase (FDC1)
3-Octaprenyl-4-hydroxybenzoate decarboxylase (OhbA1)
cis-Aconitate decarboxylase (CadA1)
Mevalonate diphosphate decarboxylase (*P. torridus* MVD)
Mevalonate diphosphate decarboxylase (*S. cerevisiae* MVD)
α-Amino-*β*-carboxymuconate-*ε*-semialdehyde decarboxylase (ACMSD)

- Each enzyme was successfully isolated and expressed in *E. coli*.
- After confirmation of enzyme activity on the natural substrates, the enzymes were then evaluated for the ability to form MAA. Unfortunately, no MAA was observed.
- In the attempt of generating the desired activity, a focused mutagenesis-based protein engineering approach was performed on two of the selected enzymes, the MVD from *S. cerevisiae* and from *P. torridus*.

- Upon successful generation of libraries, enzyme variants were screened using a high throughput screening protocol developed in the course of the project, based on the monitoring of gaseous CO_2 in 96-well plate format. Unfortunately, MAA formation was not achieved.
- Previously unknown information regarding some of the candidate enzymes were unveiled
  - The role of the two enzymes involved in decarboxylation of phenylacrylic acids in *S. cerevisiae* and *A. niger* was partially clarified. FDC1 and OhbA1, and not PAD1 and PadA1, are the main catalysts involved in decarboxylation. OhbA1 requires the presence of a novel organic cofactor, and a possible role of PadA1 in providing these cofactors is possible[1].
  - The previously suggested catalytic mechanism of ACMSD was questioned. Different to what believed, it was demonstrated that the extra moieties present on the natural substrate, and absent on the tested non-natural ones, may be directly involved in catalysis. Moreover, further investigation revealed previously unknown compounds that function as inhibitors.
  - *P. torridus* MVD was discovered to belong to a new class of enzymes, mevalonate-3-kinase, and not to MVDs, and it was able to catalyse the formation of isobutene from 3-HIV with unprecedented rates. M3K is part of a novel MVA pathway recently discovered in Archaea.

[1] Just before submission of this thesis, two papers were published proposing that padA1 is not a decarboxylase *per se* but a flavin prenyl transferase that forms a prenylated flavin (prFMN). prFMN is required by OhbA1 for decarboxylation which occurs *via* 1,3 dipolar cycloaddition (Payne *et al.*, 2015; White *et al.*, 2015).This study reveals the functional relationship between the two enzymes

6 Discussion and future work

During the course of this project, priority was given in using mevalonate diphosphate decarboxylases (MVDs). These enzymes could have potentially catalysed the formation of MAA from citramalic acid, metabolic intermediate for which an efficient bioproduction process from renewable sources have now been developed (Eastham *et al.*, 2015). The wild type enzymes were not able to recognise citramalic acid and focused modifications of the binding pocket to develop substrate acceptance were unsuccessful. An alternative approach would be to evaluate distal mutations, as they can also highly affect substrate biding and catalytic properties of an enzyme (Wong *et al.*, 2005). This evaluation was not performed due to the limited throughput level of the screening assay developed in the course of this project. The evaluation of distal mutations is usually performed upon random introduction of variation along the whole gene sequence and upon screening of a large number of enzymatic variants. This could not be done with the Microresp^{TM}-based assay.

Therefore, in order to test the effect of distal mutations on MVDs, an alternative screening method is necessary. Future work should be focused in developing such a screening. An interesting approach would be to use *E. coli* engineered for citramalic acid production (Eastham *et al.*, 2015) to develop an *in vivo* selection strategy based on MAA toxicity. Citramalic acid is significantly less toxic than MAA to the cells (Eastham *et al.*, 2015) and, if converted to MAA through an active MVD, the cells will likely not grow or grow slower. MVD variants can be generated and expressed in the citramalic acid producing *E. coli* to exploit this characteristic. A limitation of this approach is that the amount of MAA produced must be high enough to develop the toxic effect and, if the MVD is not highly efficient, this target may not be achieved.

In this thesis, concerns regarding the fact that the tested MVDs were not active on citramalic acid because of the presence of a charged carboxylate, absent on the natural substrate of the enzyme, which may interfere with the catalytic residues, were expressed. In order to confirm this suggestion, the enzymes could be tested on a citramalic acid derivative where the carboxylate is shielded, for example, with a methyl or ethyl ester. Since methyl or ethyl citramalic acid are very hard to obtain by chemical synthesis and cannot be isolated from biological sources, this investigation was not performed. A possible chemo-biological route to obtain ethyl citramalate was initially considered but then not tested for lack of time (Figure 6.1).

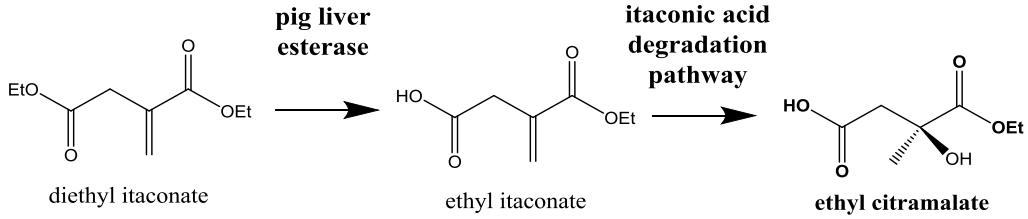


Figure 6.1: Potential route for production of ethyl citramalate

The route starts from diethyl itaconate, easily produced through esterification of itaconic acid with ethanol, which is then converted to the monoester by digestion with pig liver esterase (Ferraboschi *et al.*, 1994). The conversion of ethyl itaconate to ethyl citramalate could potentially be achieved using the itaconic acid degradation pathway discovered in *Alcaligenes sp.* and *Pseoudomonas sp.* (Bing fang *et al.*, 2000; Cooper and Kornberg, 1964; Ueda *et al.*, 1993), a route that normally converts itaconic acid to (*S*)-citramalic acid (Figure 2.9, Page 13). In order for this route to be possible, the enzymes involved in the pathway need to be able to act on the ethyl derivative as well, and this could present a significant limitation. Nevertheless, this route provides a new way of producing ethyl citramalate and would consent to test this substrate for decarboxylation by MVDs.

The putative MVD from *P. torridus* was discovered to be a mevalonate-3-kinase (M3K). However, when performing protein engineering, this enzyme was processed as an MVD, and libraries were designed on the base of a homology model built on MVD templates (see section 5.5.2.2, Page 123). The nature of the enzyme has now been clarified and the structure of the homologue from *T. acidophilum*, co-crystallised with its natural substrate, have been very recently unveiled (Vinokur *et al.*, 2015). Therefore, there is now the possibility to design more precise libraries by generating a new, more accurate model of *P. torridus* M3K using *T. acidophilum* M3K as a template. Although M3K is a kinase and not a decarboxylase *per se*, successful phosphorylation of citramalic acid would likely generate an unstable quaternary carbon intermediate that spontaneously dephosphorylates/decarboxylates to MAA. Experimental evidence of this was confidentially supplied by Ingenza Ltd (Edinburgh, UK). Therefore, this could still present an attractive route for the bioproduction of MAA. Investigation of the use of other kinases (EC 2.7), and in particular among those able to transfer phosphates to alcohol groups as acceptors (EC 2.7.1), is also an attractive option to explore. The EC 2.7.1 class of enzymes is very broad and, to date, around 280 sub-classes (including M3K discovered in this work) have already been identified (MetaCyc, 2015). Therefore, there is significant scope to find an active catalyst for the phosphorylation of citramalic acid.

Citramalic acid could also potentially be converted to a sulphate derivative, where a sulphate group rather than a phosphate is transferred to the hydroxyl group on the C3 the substrate. The sulphate citramalic acid, as the phosphate counterpart, should also be equally unstable and undergo spontaneous desulphonation/decarboxylation. Enzymes of the class of sulphotransferases (ST; EC 2.8) are able to catalyse such a reaction and, therefore, there is significant scope for finding active catalysts also in this class. Some STs have been already tested in a patent for the production of different alkenes through sulphorylation of

3-hydroxy alkenoic acids (Marliere *et al.*, 2014), but details were not disclosed.

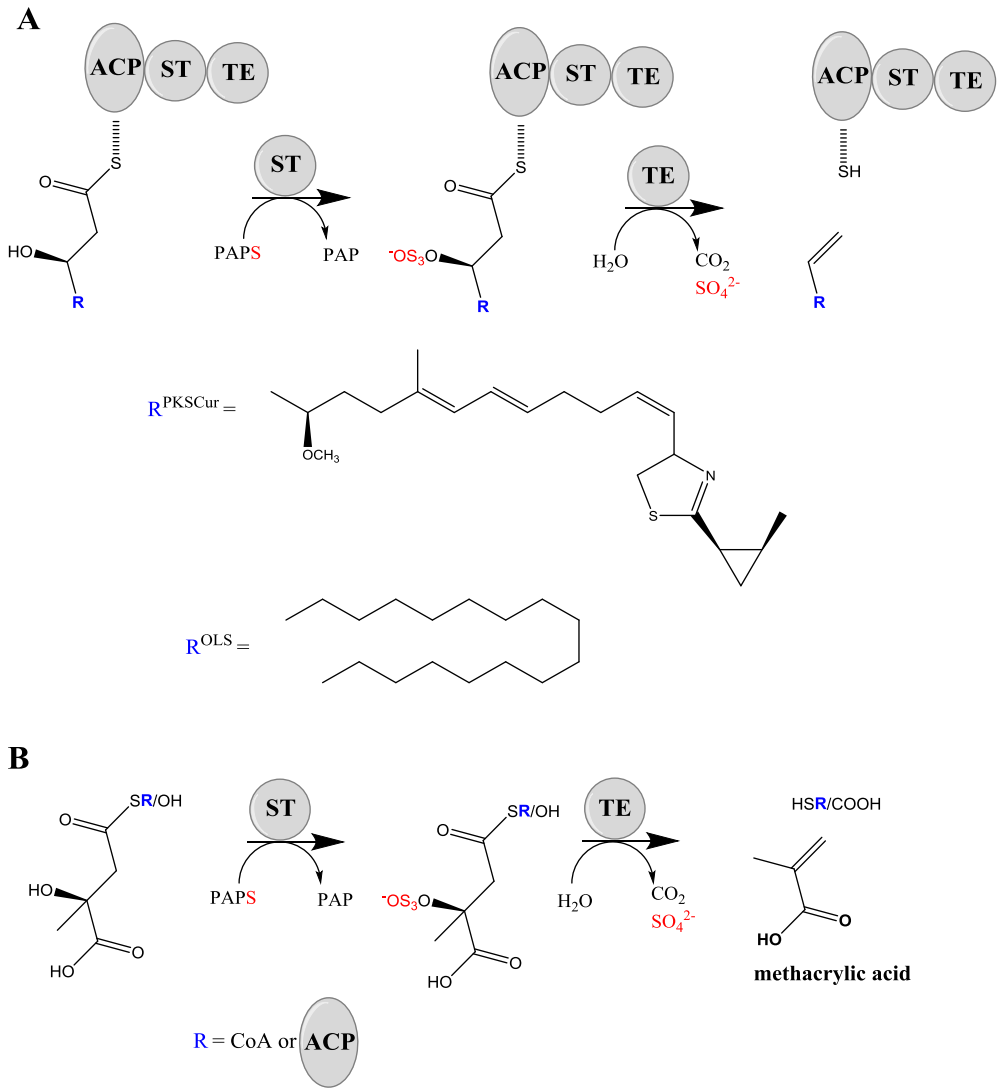


Figure 6.2: Formation of terminal alkenes through sulfonylation/decarboxylation (A) The sulfotransferase domain (ST) and the thioesterase domain (TE) of the CurM polyketide synthase of the curacin A biosynthetic pathway found in *Moorea producens* (PKSCur) and of the olefin synthase system found in *Synechococcus* PCC 7002 (OLS) catalyse the formation of a terminal alkene from a *β*-hydroxy acyl thioester substrate, usually linked to an acyl carrier (ACP). Phosphoadenosine 5'-phosphosulfate (PAPS) is required as sulfonate donor. (B) ST and TE could potentially catalyse the sulfonylation/decarboxylation of citramalic acid derivatives to form methacrylic acid (MAA).

Of particular interest is the ST domain of the CurM polyketide synthase of the curacin A biosynthetic pathway found in *Moorea producens* (PKSCur) (Gehret *et al.*, 2011; Gu *et al.*, 2006; Gu *et al.*, 2009; McCarthy *et al.*, 2012) and of the olefin synthase system found in *Synechococcus* PCC 7002 (OLS) (Mendez-Perez *et al.*, 2011). In both cases, the ST domain is part of a larger

multifunctional polypeptide linked to an acyl carrier protein domain (ACP) and to a thioesterase domain (TE) (Figure 6.2A). The ST domain catalyses the sulphonylation of a *β*-hydroxy acyl thioester substrate, and the sulphate substrate is subsequently decarboxylated and desulphonated by the TE domain, which releases a terminal alkene. Interestingly, although the substrate must be linked to the ACP domain to reach the maximum efficiency of the reaction, activity was also observed when testing the CoA derivatives and the free acids (Gu *et al.*, 2006).

The utilisation of this ST/TE system for the conversion of ACP-citramalic acid, citramalyl-CoA or citramalic acid to MAA *via* sulphate citramalate (Figure 6.2B) might be an attractive route for future research. However, the presence of the methyl substituent and, especially the carboxylate, present on citramalic acid but absent on the natural substrates, could potentially still present an issue (Figure 6.2). Various synthetic substrates have been successfully decarboxylated by these enzymes but none presents a carboxylate moiety so close to the hydroxyl of the substrate (Gu *et al.*, 2006; McCarthy *et al.*, 2012). Protein engineering could potentially be attempted to overcome the problem and evolve substrate acceptance. Structural and functional data on both the ST and TE domains have been recently reported (Gehret *et al.*, 2011; McCarthy *et al.*, 2012), and could be used to design an accurate mutagenesis strategy.

In this work, only a few decarboxylases were investigated, carefully chosen based on mechanistic and functional characteristics that were believed to be important for the target reactions. Following the suggestions made in this work, it would be possible to screen other decarboxylases with slightly different properties. The substrates of interest (itaconic, mesaconic, citraconic and citramalic acid) are all dicarboxylic acids. From this study, it is clear that the presence of such decarboxylate moieties can prevent enzymes from binding/acting with the non-natural substrates. Therefore, new candidates should be sought amongst those decarboxylases which

already have developed binding properties towards dicarboxylic acids. Unfortunately, to date, among the decarboxylases with this feature, there is none that raises particular interest. Some amino acid decarboxylases, such as aspartate (EC 4.1.1.11) or glutamate (EC 4.1.1.15) decarboxylase, are able to accept dicarboxylate moieties. However, decarboxylation occurs through the formation of an amino acid-pyridoxal phosphate or -pyruvoyl Schiff-base that requires the presence of an amino moiety on the substrate (Toney, 2005; Webb and Abell, 2003), something that itaconic, mesaconic, citraconic or citramalic acid lack. In fact, citraconic and mesaconic acid have already been shown to inhibit glutamate decarboxylase and the enzyme was not able to catalyse their decarboxylation (Fonda, 1972). Alternatively oxaloacetate decarboxylase (OAD; EC 4.1.1.3) and malonate decarboxylase (MD; EC 4.1.1.88 and 4.1.1.89) are able to bind compounds with two carboxylates. However, the overall chemical structure oxaloacetic and malonic acid differs significantly from that of itaconic, mesaconic, citraconic and citramalic acid (Figure 6.3). Both oxaloacetate and malonate lack the double bond characteristic of itaconic, mesaconic and citraconic acid. The carbonyl oxygen of oxaloacetate, missing on the acids of interest, was suggested to be necessary for the reaction catalysed by OAD (Dimroth *et al.*, 2001).

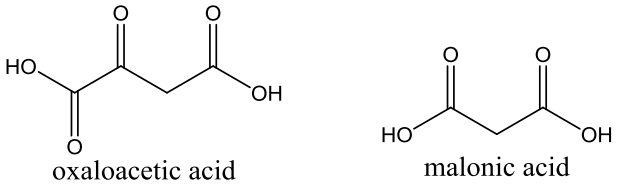


Figure 6.3: Structure of oxaloacetic and malonic acid

Nevertheless, the range of decarboxylases available in nature is broad. Some of them are still poorly characterised and some others have just been discovered (Li *et al.*, 2012b). Alternative biochemical and mechanistic properties keep emerging and it will be important to keep an eye open for when enzymes with attractive characteristics will be unveiled.

Beyond MAA formation, this study provides insights for future work in other areas. For example, it was demonstrated that the activity of phenyl acrylic acid decarboxylases (PADs) is significantly enhanced by the presence of Mg^{2+}, that it may require the presence of an organic cofactor, and that there may be the need of additional enzymes to produce it and to recycle it. This information can now be taken into consideration when using PADs in the bioproduction of styrene and other monounsaturated alkenes, and better metabolic engineering strategies can be designed to enhance the existing processes which are still characterised by poor yields and efficiencies (Leys *et al.*, 2013; McKenna and Nielsen, 2011; McKenna *et al.*, 2014).

Also, new inhibitors of α-amino-β-carboxymuconate-ε-semialdehyde decarboxylase (ACMSD) were unveiled. Two of this inhibitors, mesaconic and fumaric acid, present a dicarboxylic acid moiety in *trans* configuration, uncommon compared to classical inhibitors, which have the moiety in *trans* configuration instead (Fukuwatari *et al.*, 2004a; Fukuwatari *et al.*, 2004b; Fukuwatari *et al.*, 2002; Huo *et al.*, 2013; Shibata *et al.*, 2007). Since this enzyme is involved in the development of numerous diseases in humans (Guidetti and Schwarcz, 2003; Guillemin *et al.*, 2003; Mehler *et al.*, 1964; Reinhard, 2004; Schwarcz, 2004), the design of new inihbitors which may provide better drugs and treatments is of extreme important, and this study provides further insights into this context.

Finally, the mutant libraries generated for *S. cerevisiae* MVD and *P. torridus* M3K could be used for purposes that go beyond MAA formation. These libraries were only screened for activity on citramalic acid but, potentially, could provide improved enzymes for the formation of isobutene from 3-hydroxyisovaleric acid, already reported with MVDs (Gogerty and Bobik, 2010; Marliere *et al.*, 2011). Morevoer, as both libraries were designed to modify the binding pocket and to widen the substrate range of the enzymes, they could also be tested on several other substrates in the attempt

to produce other terminal alkenes. For examples, they could be screened for formation of ethene and propene from 3-hydroxybutyrate and 3-hyroxypropionate, respectively, reactions not observed with the wild type enzymes.

7 Conclusions

Decarboxylases are a very useful tool for the defunctionalisation of organic molecules and have been extensively applied for the synthesis of chemicals in the past (Kourist *et al.*, 2014).This work provides the first investigation on the use of decarboxylases for the bioproduction of MAA, and lays the foundation for the future research in this direction. Although, at present, the approach does not seem feasible, the findings reported contribute in expanding the knowledge around the use of these enzymes in biocatalysis. Novel insights on previously poorly characterised decarboxylases were unveiled, including newly proposed enzymatic mechanisms and cofactor requirements. The information obtained can be used to perform more focused research on what is needed for production of MAA, as well as to fine tune the search for other suitable candidates. Moreover, since some of the tested enzymes have also demonstrated potential for the bioproduction of other useful chemicals, this study provides information that could be beneficial to further enhance these processes too. Finally, beyond MAA formation and biocatalysis, this study reveals novel and interesting insights on some of the investigated catalysts, such as the discovery of a new class of enzyme and the finding of new inhibitors of an important molecular target for several human diseases.

8 References

ALEXANDER, F. W., SANDMEIER, E., MEHTA, P. K. & CHRISTEN, P. 1994. Evolutionary relationships among pyridoxal-5'-phosphate-dependent enzymes. *Eur J Biochem,* 219**,** 953-960.

ALONSO-GUTIERREZ, J., CHAN, R., BATTH, T. S., ADAMS, P. D., KEASLING, J. D., PETZOLD, C. J. & LEE, T. S. 2013. Metabolic engineering of *Escherichia coli* for limonene and perillyl alcohol production. *Metab Eng,* 19**,** 33-41.

ALTSCHUL, S. F., MADDEN, T. L., SCHAFFER, A. A., ZHANG, J., ZHANG, Z., MILLER, W. & LIPMAN, D. J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res,* 25**,** 3389-3402.

ALVEAR, M., JABALQUINTO, A. M., EYZAGUIRRE, J. & CARDEMIL, E. 1982. Purification and characterization of avian liver mevalonate-5-pyrophosphate decarboxylase. *Biochemistry,* 21**,** 4646-4650.

AMIN, A. S., DESSOUKI, H. A. & KHALIL, K. M. 2010. Indirect spectrophotometric determination of piroxicam and tenoxicam through oxidation with potassium permanganate. *Bull Chem Soc Ethiop,* 24**,** 121-126.

ANTONELLI, A. & FABBRI, C. 1998. Study on Roman chamomile (Chamaemelum nobile L. All.) oil. *Journal of Essential Oil Research,* 10**,** 571-574.

ARNOLD, F. H. 1996. Directed evolution: Creating biocatalysts for the future. *Chemical Engineering Science,* 51**,** 5091-5102.

ARNOLD, F. H. & GEORGIOU, G. 2003. *Directed enzyme evolution: screening and selection methods, Vol. 230*, Humana, Totowa.

ATSUMI, S. & LIAO, J. C. 2008. Directed evolution of *Methanococcus jannaschii* citramalate synthase for biosynthesis of 1-propanol and 1-butanol by *Escherichia coli*. *Appl Environ Microbiol,* 74**,** 7802-7808.

AVILA, C., KORNILAYEV, B. A. & BLAGG, B. S. 2006. Development and optimization of a useful assay for determining Hsp90's inherent ATPase activity. *Bioorg Med Chem,* 14**,** 1134-1142.

AZAMI, Y., HATTORI, A., NISHIMURA, H., KAWAIDE, H., YOSHIMURA, T. & HEMMI, H. 2014. (*R*)-Mevalonate 3-phosphate is an intermediate of the mevalonate pathway in *Thermoplasma acidophilum*. *J Biol Chem,* 289**,** 15957-15967.

BAR-PELED, M., GRIFFITH, C. L. & DOERING, T. L. 2001. Functional cloning and characterization of a UDP- glucuronic acid decarboxylase: The pathogenic fungus *Cryptococcus neoformans* elucidates UDP-xylose synthesis. *Proc Natl Acad Sci,* 98**,** 12003-12008.

BARTA, M. L., SKAFF, D. A., MCWHORTER, W. J., HERDENDORF, T. J., MIZIORKO, H. M. & GEISBRECHT, B. V. 2011. Crystal structures of *Staphylococcus epidermidis* mevalonate diphosphate decarboxylase bound to inhibitory

analogs reveal new insight into substrate binding and catalysis. *J Biol Chem,* 286**,** 23900-23910.

BARTA, M. L., MCWHORTER, W. J., MIZIORKO, H. M. & GEISBRECHT, B. V. 2012. Structural basis for nucleotide binding and reaction catalysis in mevalonate diphosphate decarboxylase. *Biochemistry,* 51**,** 5611-5621.

BARTSCH, S., KOURIST, R. & BORNSCHEUER, U. T. 2008. Complete inversion of enantioselectivity towards acetylated tertiary alcohols by a double mutant of a *Bacillus subtilis* esterase. *Angew Chem Int Ed Engl,* 47**,** 1508-1511.

BAUER, W. 2000. Methacrylic Acid and Derivatives. *Ullmann's Encyclopedia of Industrial Chemistry.* Wiley-VCH Verlag GmbH & Co. KGaA.

BAUMANN, M., STÜRMER, R. & BORNSCHEUER, U. T. 2001. A high-throughput-screening method for the identification of active and enantioselective hydrolases. *Angew Chem Int Ed Engl,* 40**,** 4201-4204.

BECKER, J. & WITTMANN, C. 2015. Advanced biotechnology: metabolically engineered cells for the bio-based production of chemicals and fuels, materials, and health-care products. *Angew Chem Int Ed Engl,* 54**,** 3328-50.

BENSON, D. A., CAVANAUGH, M., CLARK, K., KARSCH-MIZRACHI, I., LIPMAN, D. J., OSTELL, J. & SAYERS, E. W. 2013. GenBank. *Nucleic Acids Res,* 41**,** D36-D42.

BENTLEY, R. & THIESSEN, C. P. 1957. Biosynthesis of itaconic acid in *A. terreus*. III. The properties and reaction mechanism of *cis*-aconitic acid decarboxylase. *J Biol Chem,* 226**,** 703-720.

BERG, I. A. & IVANOVSKY, R. N. 2009. Enzymes of the citramalate cycle in *Rhodospirillum rubrum*. *Microbiology,* 78**,** 16-24.

BING FANG, H., TETSUO, O., TOSHIAKI NAKAJIMA, K. & TADAATSU, N. 2000. Efficient conversion of itaconic acid to (*s*)-(+)-citramalic acid by *Alcaligenes xylosoxydans* IL42. *J Biosci Bioeng,* 89**,** 388-391.

BLAZECK, J., MILLER, J., PAN, A., GENGLER, J., HOLDEN, C., JAMOUSSI, M. & ALPER, H. S. 2014. Metabolic engineering of *Saccharomyces cerevisiae* for itaconic acid production. *Appl Microbiol Biotechnol,* 98**,** 8155-8164.

BLOCH, K. 1965. The biological synthesis of cholesterol. *Science* 150**,** 19-28.

BLUMHOFF, M. L., STEIGER, M. G., MATTANOVICH, D. & SAUER, M. 2013. Targeting enzymes to the right compartment: metabolic engineering for itaconic acid production by *Aspergillus niger*. *Metab Eng,* 29**,** 26-32.

BOERSMA, Y. L., DRÖGE, M. J., VAN DER SLOOT, A. M., PIJNING, T., COOL, R. H., DIJKSTRA, B. W. & QUAX, W. J. 2008. A novel genetic selection system for improved enantioselectivity of *Bacillus subtilis* Lipase A. *ChemBioChem,* 9**,** 1110-1115.

BONAGURA, C. A., BHASKAR, B., SUNDARAMOORTHY, M. & POULOS, T. L. 1999. Conversion of an engineered potassium-binding site into a calcium-selective site in cytochrome c peroxidase. *J Biol Chem,* 274**,** 37827-37833.

BONANNO, J. B., EDO, C., ESWAR, N., PIEPER, U., ROMANOWSKI, M. J., ILYIN, V., GERCHMAN, S. E., KYCIA, H., STUDIER, F. W., SALI, A. & BURLEY, S. K. 2001. Structural genomics of enzymes involved in sterol/isoprenoid biosynthesis. *Proc Natl Acad Sci U S A,* 98**,** 12896-12901.

BONNARME, P., GILLET, B., SEPULCHRE, A. M., ROLE, C., BELOEIL, J. C. & DUCROCQ, C. 1995. Itaconate biosynthesis in *Aspergillus terreus*. *J Bacteriol,* 177**,** 3573-3578.

BORNSCHEUER, U. T. & KAZLAUSKAS, R. J. 2004. Catalytic promiscuity in biocatalysis: Using old enzymes to form new bonds and follow new pathways. *Angew Chem Int Ed Engl,* 43**,** 6032-6040.

BORNSCHEUER, U. T., HUISMAN, G. W., KAZLAUSKAS, R. J., LUTZ, S., MOORE, J. C. & ROBINS, K. 2012. Engineering the third wave of biocatalysis. *Nature,* 485**,** 185-194.

BOS, C. J., DEBETS, A. J. M., SWART, K., HUYBERS, A., KOBUS, G. & SLAKHORST, S. M. 1988. Genetic analysis and the construction of master strains for assignment of genes to six linkage groups in *Aspergillus niger*. *Curr Genet,* 14**,** 437-443.

BOSLEY, A. D. & OSTERMEIER, M. 2005. Mathematical expressions useful in the construction, description and evaluation of protein libraries. *Biomol Eng,* 22**,** 57-61.

BOUGIOUKOU, D. J., KILLE, S., TAGLIEBER, A. & REETZ, M. T. 2009. Directed evolution of an enantioselective enoate-reductase: Testing the utility of iterative saturation mutagenesis. *Adv Synth Catal,* 351**,** 3287-3305.

BRANNIGAN, J. A. & WILKINSON, A. J. 2002. Protein engineering 20 years on. *Nat Rev Mol Cell Biol,* 3**,** 964-970.

BROOKSBANK, C., BERGMAN, M. T., APWEILER, R., BIRNEY, E. & THORNTON, J. 2014. The European Bioinformatics Institute's data resources 2014. *Nucleic Acids Res,* 42**,** D18-D25.

BROWN, C., LEIJON, F. & BULONE, V. 2012. Radiometric and spectrophotometric in vitro assays of glycosyltransferases involved in plant cell wall carbohydrate biosynthesis. *Nat Protoc,* 7**,** 1634-1650.

BUCKEL, W. 2001. Sodium ion-translocating decarboxylases. *Biochim Biophys Acta,* 1505**,** 15-27.

BYRES, E., MARTIN, D. M. A. & HUNTER, W. N. 2005. A preliminary crystallographic analysis of the putative mevalonate diphosphate decarboxylase from*Trypanosoma brucei*. *Acta Crystallogr F Struct Biol Commun.,* 61**,** 581-584.

BYRES, E., ALPHEY, M. S., SMITH, T. K. & HUNTER, W. N. 2007. Crystal Structures of *Trypanosoma brucei* and *Staphylococcus aureus* mevalonate diphosphate decarboxylase inform on the determinants of specificity and reactivity. *J Mol Biol,* 371**,** 540-553.

CAMPBELL, C. D., CHAPMAN, S. J., CAMERON, C. M., DAVIDSON, M. S. & POTTS, J. M. 2003. A rapid microtiter plate method to measure carbon dioxide evolved from carbon substrate amendments so as to determine the physiological profiles of soil microbial communities by using whole soil. *Appl Environ Microbiol,* 69**,** 3593-3599.

CEDRONE, F., MÉNEZ, A. & QUÉMÉNEUR, E. 2000. Tailoring new enzyme functions by rational redesign. *Current Opinion in Structural Biology,* 10**,** 405-410.

CHAPMAN, E. & WONG, C.-H. 2002. A pH sensitive colorometric assay for the high-throughput screening of enzyme inhibitors and substrates : A case study using kinases. *Bioorg Med Chem,* 10**,** 551-555.

CHEN , H., BØRJESSON, U., ENGKVIST, O., KOGEJ, T., SVENSSON , M. A., BLOMBERG, N., WEIGELT, D., BURROWS, J. N. & LANGE, T. 2009. ProSAR: A new methodology for combinatorial library design. *J Chem Inf Model,* 49**,** 603-614.

CHEN, K. & ARNOLD, F. H. 1993. Tuning the activity of an enzyme for unusual environments: sequential random mutagenesis of subtilisin E for catalysis in dimethylformamide. *Proc Natl Acad Sci U S A,* 90**,** 5618-5622.

CHEN, M. & POULTER, C. D. 2010. Characterization of thermophilic archaeal isopentenyl phosphate kinases. *Biochemistry,* 49**,** 207-217.

CHICA, R. A., DOUCET, N. & PELLETIER, J. N. 2005. Semi-rational approaches to engineering enzyme activity: combining the benefits of directed evolution and rational design. *Curr Opin Biotechnol,* 16**,** 378-384.

CHITRA, K., SUJATHA, K., AHMED, I. R., SHALINI, K. & PRIYA, B. L. 2004. Spectrophotometric estimation of tranexamic acid in bulk and pharmaceutical dosage form. *Indian J Pharm Sci,* 66**,** 836-839.

CHUN, G.-T. & KIM, S. J. Strain improvement and statistical medium optimization for enhanced production of itaconic acid by the recombinant PD mutants of *Aspergillus terreus*. 2014. American Chemical Society, BIOT-203.

CLAUSEN, M., LAMB, C. J., MEGNET, R. & DOERNER, P. W. 1994. PAD1 encodes phenylacrylic acid decarboxylase which confers resistance to cinnamic acid in *Saccharomyces cerevisiae*. *Gene,* 142**,** 107-112.

COCO, W. M., LEVINSON, W. E., CRIST, M. J., HEKTOR, H. J., DARZINS, A., PIENKOS, P. T., SQUIRES, C. H. & MONTICELLO, D. J. 2001. DNA shuffling method for generating highly recombined genes and evolved enzymes. *Nat Biotechnol,* 19**,** 354-359.

COLE, M. F. & GAUCHER, E. A. 2011. Utilizing natural diversity to evolve protein function: applications towards thermostability. *Curr Opin Chem Biol,* 15**,** 399-406.

CONTI, G., POLLEGIONI, L., MOLLA, G. & ROSINI, E. 2014. Strategic manipulation of an industrial biocatalyst - evolution of a cephalosporin C acylase. *FEBS J,* 281**,** 2443-2455.

COOPER, R. A. & KORNBERG, H. L. 1964. The utilization of itaconate by *Pseudomonas sp*. *Biochem J,* 91**,** 82-91.

CORDIER, H., KARST, F. & BERGÈS, T. 1999a. Heterologous expression in *Saccharomyces cerevisiae* of an *Arabidopsis thaliana* cDNA encoding mevalonate diphosphate decarboxylase. *Plant Mol Biol,* 39**,** 953-967.

CORDIER, H., LACOMBE, C., KARST, F. & BERGÈS, T. 1999b. The *Saccharomyces cerevisiae* mevalonate diphosphate decarboxylase (*Erg19p*) forms homodimers in vivo, and a single substitution in a structurally conserved region impairs dimerization. *Curr Microbiol,* 38**,** 290-294.

COWARD-KELLY, G. & CHEN, R. R. 2007. A window into biocatalysis and biotransformation. *Biotechnol Prog,* 23**,** 52-54.

CRAMERI, A., CWIRLA, S. & STEMMER, W. P. C. 1996a. Construction and evolution of antibody-phage libraries by DNA shuffling. *Nat Med,* 2**,** 100-102.

CRAMERI, A., WHITEHORN, E. A., TATE, E. & STEMMER, W. P. C. 1996b. Improved green fluorescent protein by molecular evolution using DNA shuffling. *Nat Biotech,* 14**,** 315-319.

CUIA, W., DEWITTA, J. G., MILLERB, S. M. & WUA, W. 1999. No metal cofactor in orotidine 5′-monophosphate decarboxylase. *Biochem Biophys Res Commun,* 259**,** 133-135.

DAHL, R. H., ZHANG, F., ALONSO-GUTIERREZ, J., BAIDOO, E., BATTH, T. S., REDDING-JOHANSON, A. M., PETZOLD, C. J., MUKHOPADHYAY, A., LEE, T. S., ADAMS, P. D. & KEASLING, J. D. 2013. Engineering dynamic pathway regulation using stress-response promoters. *Nat Biotechnol,* 31**,** 1039-1046.

DASSANAYAKE R., CAO L., SAMARANAYAKE L. & BERGES L. 2002. Characterization, heterologous expression and functional analysis of mevalonate diphosphate decarboxylase gene (MVD) of *Candida albicans*. *Mol Genet Genomics,* 267**,** 281-290.

DAVIDS, T., SCHMIDT, M., BOTTCHER, D. & BORNSCHEUER, U. T. 2013. Strategies for the discovery and engineering of enzymes for biocatalysis. *Curr Opin Chem Biol,* 17**,** 215-220.

DE GUZMAN, D. 2011. *Rubber industry seeks bio-based chemicals potential* [Online]. Available: http://www.icis.com/Articles/2011/07/01/9481223/Rubber-industry-seeks-bio-based-chemicals-potential.html [Accessed March 2015].

DE GUZMAN, D. 2013. *Bio-MMA development expands* [Online]. Available: http://greenchemicalsblog.com/2013/03/14/bio-mma-development-expands/ [Accessed March 2015].

DELCOURT, M., ANISSIMOVA, M. & MARLIERE, P. 2013. *Enzymatic production of isoprenol from mevalonate by diphosphomevalonate decarboxylases*. WO2013150100A1.

DELLAS, N., THOMAS, S. T., MANNING, G. & NOEL, J. P. 2013. Discovery of a metabolic alternative to the classical mevalonate pathway. *Elife,* 2**,** e00672.

DHE-PAGANON, S., MAGRATHV J & RH., A. 1994. Mechanism of mevalonate pyrophosphate decarboxylase: evidence for a carbocationic transition state. *Biochemistry,* 33**,** 13355-13362.

DIMROTH, P., JOCKEL, P. & SCHMID, M. 2001. Coupling mechanism of the oxaloacetate decarboxylase Na+ pump. *Biochim Biophys Acta,* 1505**,** 1-14.

DRAGE, S., ENGELMEIER, D., BACHMANN, G., SESSITSCH, A., MITTER, B. & HADACEK, F. 2012. Combining microdilution with MicroResp: microbial substrate utilization, antimicrobial susceptibility and respiration. *J Microbiol Methods,* 88**,** 399-412.

DWIARTI, L., YAMANE, K., YAMATANI, H., KAHAR, P. & OKABE, M. 2002. Purification and characterization of cis-aconitic acid decarboxylase from *Aspergillus terreus* TN484-M1. *J Biosci Bioeng,* 94**,** 29-33.

EASTHAM, G. R., JOHNSON, D. W., STRAATHOF, A. J. J., FRAAIJE, M. W. & WINTER, R. T. 2013. *Chemoenzymic process for the production of methyl methacrylate*. WO2013179005A1.

EASTHAM, G. R., JOHNSON, D. W., ARCHER, I., CARR, R., WEBB, J. & STEPHENS, G. 2015. *A process for production of methacrylic acid and derivatives thereof*. WO2015022496A2.

EDGAR, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res,* 32**,** 1792-1797.

EGGERT, T., REETZ, M. T. & JAEGER , K.-E. 2004. *Enzyme functionality-design, engineering, and screening (Ed. A. Svendsen)*, Marcel Dekker, New York.

FENG, X., SANCHIS, J., REETZ, M. T. & RABITZ, H. 2012. Enhancing the efficiency of directed evolution in focused enzyme libraries by the adaptive substituent reordering algorithm. *Chemistry,* 18**,** 5646-5654.

FERRABOSCHI, P., CASATI, S., GRISENTI, P. & SANTANIELLO, E. 1994. Selective enzymatic transformations of itaconic acid derivatives: An access to potentially useful building blocks. *Tetrahedron,* 50**,** 3251-3258.

FIKHTENGOL'TS, V. S., ZOLOTAREVA, R. V. & L'VOV, Y. A. 1969. *Atlas of ultraviolet absorption spectra of substances used in producing synthetic rubber. 2nd ed*, Khimiya, Leningrad. Otd.

FOLKERTSMA, S., VAN NOORT, P., VAN DURME, J., JOOSTEN, H. J., BETTLER, E., FLEUREN, W., OLIVEIRA, L., HORN, F., DE VLIEG, J. & VRIEND, G. 2004. A family-based approach reveals the function of residues in the nuclear receptor ligand-binding domain. *J Mol Biol,* 341**,** 321-335.

FONDA, M. L. 1972. Glutamate decarboxylase. Substrate specificity and inhibition by carboxylic acids. *Biochemistry,* 11**,** 1304-1309.

FUJIHASHI, M., ISHIDA, T., KURODA, S., KOTRA, L. P., PAI, E. F. & MIKI, K. 2013. Substrate distortion contributes to the catalysis of orotidine 5'-monophosphate decarboxylase. *J Am Chem Soc,* 135**,** 17432-43.

FUKUWATARI, T., SUZUKI, Y., SUGIMOTO, E. & SHIBATA, K. 2002. Identification of a toxic mechanism of the plasticizers, phtahlic acid esters, which are putative endocrine disrupters: time-dependent increase in quinolinic acid and its metabolites in rats fed di(2-ethylhexyl)phthalate. *Biosci Biotechnol Biochem,* 66**,** 2687-2691.

FUKUWATARI, T., OHSAKI, S., FUKUOKA, S., SASAKI, R. & SHIBATA, K. 2004a. Phthalate esters enhance quinolinate production by inhibiting alpha-amino-beta-carboxymuconate-epsilon-semialdehyde decarboxylase (ACMSD), a key enzyme of the tryptophan pathway. *Toxicol Sci,* 81**,** 302-308.

FUKUWATARI, T., OHTA, M., SUGIMOTO, E., SASAKI, R. & SHIBATA, K. 2004b. Effects of dietary di(2-ethylhexyl)phthalate, a putative endocrine disrupter, on enzyme activities involved in the metabolism of tryptophan to niacin in rats. *Biochim Biophys Acta,* 1672**,** 67-75.

GARAVAGLIA, S., PEROZZI, S., GALEAZZI, L., RAFFAELLI, N. & RIZZI, M. 2009. The crystal structure of human alpha-amino-beta-carboxymuconate-epsilon-semialdehyde decarboxylase in complex with 1,3-dihydroxyacetonephosphate suggests a regulatory link between NAD synthesis and glycolysis. *FEBS J,* 276**,** 6615-23.

GARCÍA-PALACIOS, P., MILLA, R., ÁLVARO-SÁNCHEZ, M., MARTÍN-ROBLES, N. & MAESTRO, M. 2013. Application of a high-throughput laboratory method to assess litter decomposition rates in multiple-species experiments. *Soil Biol Biochem,* 57**,** 929-932.

GASTEIGER, E., GATTIKER, A., HOOGLAND, C., IVANYI, I., APPEL, R. D. & BAIROCH, A. 2003. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res,* 31**,** 3784-3788.

GASTEIGER, E., HOOGLAND, C., GATTIKER, A., DUVAUD, S. E., WILKINS, M., APPEL, R. & BAIROCH, A. 2005. Protein identification and analysis yools on the ExPASy Server. *In:* WALKER, J. (ed.) *The Proteomics Protocols Handbook.* Humana Press.

GAWRONSKI, J. D. & BENSON, D. R. 2004. Microtiter assay for glutamine synthetase biosynthetic activity using inorganic phosphate detection. *Anal Biochem,* 327**,** 114-118.

GEHRET, J. J., GU, L., GERWICK, W. H., WIPF, P., SHERMAN, D. H. & SMITH, J. L. 2011. Terminal alkene formation by the thioesterase of curacin A biosynthesis: structure of a decarboxylating thioesterase. *J Biol Chem,* 286**,** 14445-14454.

GERSHENZON, J. & DUDAREVA, N. 2007. The function of terpene natural products in the natural world. *Nat Chem Biol,* 3**,** 408-414.

GLOBAL BIOENERGIES. 2014. *Global Bioenergies and Fraunhofer CBP take the next step towards the set-up of the Leuna industrial pilot* [Online]. Available: http://www.global-bioenergies.com/communiques/140703_pr_en.pdf [Accessed March 2015].

GODA, S. K., MINTON, N. P., BOTTING, N. P. & GANI, D. 1992. Cloning, sequencing, and expression in *Escherichia coli* of the *Clostridium*

tetanomorphum gene encoding β-methylaspartase and characterization of the recombinant protein. *Biochemistry,* 31**,** 10747-10756.

GODDARD, J. & REYMOND, J. 2004. Enzyme assays for high-throughput screening. *Curr Opin Biotechnol,* 15**,** 314-322.

GOGERTY, D. S. & BOBIK, T. A. 2010. Formation of isobutene from 3-hydroxy-3-methylbutyrate by diphosphomevalonate decarboxylase. *Appl Environ Microbiol,* 76**,** 8004-8010.

GORYANOVA, B., GOLDMAN, L. M., AMYES, T. L., GERLT, J. A. & RICHARD, J. P. 2013. Role of a Guanidinium Cation-Phosphodianion Pair in Stabilizing the Vinyl Carbanion Intermediate of Orotidine 5'-Phosphate Decarboxylase-Catalyzed Reactions. *Biochemistry,* 52**,** 7500-7511.

GOUET, P., ROBERT, X. & COURCELLE, E. 2003. ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res,* 31**,** 3320-3323.

GRISSOM, C. B. & CLELAND, W. W. 1988. Isotope effect studies of the chemical mechanism of pig heart NADP isocitrate dehydrogenase. *Biochemistry,* 27**,** 2934-2943.

GU, L., JIA, J., LIU, H., HÅKANSSON, K., GERWICK, W. H. & SHERMAN, D. H. 2006. Metabolic coupling of dehydration and decarboxylation in the curacin A pathway: Functional identification of a mechanistically diverse enzyme pair. *J Am Chem Soc,* 128**,** 9014-9015.

GU, L., WANG, B., KULKARNI, A., GEHRET, J. J., LLOYD, K. R., GERWICK, L., GERWICK, W. H., WIPF, P., HAKANSSON, K., SMITH, J. L. & SHERMAN, D. H. 2009. Polyketide decarboxylative chain termination preceded by O-sulfonation in curacin A biosynthesis. *J Am Chem Soc,* 131**,** 16033-16035.

GUIDETTI, P. & SCHWARCZ, R. 2003. 3-Hydroxykynurenine and quinolinate: pathogenic synergism in early grade Huntington's disease? *Adv Exp Med Biol,* 527**,** 137-145.

GUILLEMIN, G. J., WILLIAMS, K. R., SMITH, D. G., SMYTHE, G. A., CROITORU-LAMOURY, J. & BREW, B. J. 2003. Quinolinic acid in the pathogenesis of Alzheimer's disease. *Adv Exp Med Biol,* 527**,** 167-176.

GUPTA, D., SUMMERS, M. L. & BASU, C. 2014. Engineering an isoprenoid pathway in *Escherichia coli* for production of 2-methyl-3-buten-2-ol: a potential biofuel. *Mol Biotechnol,* 56**,** 516-523.

HARRIS, B. 2010. *Acrylics for the future* [Online]. Available: http://www.ingenia.org.uk/ingenia/issues/issue45/harris.pdf [Accessed March 2015].

HERNÁNDEZ-ACOSTA, P., SCHMID, D. G., JUNG, G., CULIÁÑEZ-MACIÀ, F. A. & KUPKE, T. 2002. Molecular characterization of the *Arabidopsis thaliana* flavoprotein AtHAL3a reveals the general reaction mechanism of 4'-phosphopantothenoylcysteine decarboxylases. *J Biol Chem,* 277**,** 20490-20498.

HILLER, K., CORDES, T. & MICHELUCCI, A. 2014. *Itaconic acid fermentation with transgenic cells expressing the mammalian IRG1 gene for aconitate decarboxylase*. WO2014161988A1.

HOLSTEIN, S. A. & HOHL, R. J. 2004. Isoprenoids: remarkable diversity of form and function. *Lipids,* 39**,** 293-309.

HORSMAN, G. P., LIU, A. M., HENKE, E., BORNSCHEUER, U. T. & KAZLAUSKAS, R. J. 2003. Mutations in distant residues moderately increase the enantioselectivity of *Pseudomonas fluorescens* esterase towards methyl 3-bromo-2-methylpropanoate and ethyl 3-phenylbutyrate. *Chemistry,* 9**,** 1933-1939.

HOWELL, D. M., XU, H. & WHITE, R. J. 1999. (*R*)-Citramalate synthase in methanogenic Archaea. *J Bacteriol,* 181**,** 331-333.

HUANG, X., LU, X., LI, Y., LI, X. & LI, J.-J. 2014. Improving itaconic acid production through genetic engineering of an industrial *Aspergillus terreus* strain. *Microb Cell Fact,* 13**,** 119.

HULLEY, M. E., TOOGOOD, H. S., FRYSZKOWSKA, A., MANSELL, D., STEPHENS, G. M., GARDINER, J. M. & SCRUTTON, N. S. 2010. Focused directed evolution of pentaerythritol tetranitrate reductase by using automated anaerobic kinetic screening of site-saturated libraries. *Chembiochem,* 11**,** 2433-2447.

HUMBLE, M. S. & BERGLUND, P. 2011. Biocatalytic Promiscuity. *European J Org Chem,* 2011**,** 3391-3401.

HUO, L., FIELDING, A. J., CHEN, Y., LI, T., IWAKI, H., HOSLER, J. P., CHEN, L., HASEGAWA, Y., QUE, L., JR. & LIU, A. 2012. Evidence for a dual role of an active site histidine in alpha-amino-beta-carboxymuconate-epsilon-semialdehyde decarboxylase. *Biochemistry,* 51**,** 5811-5821.

HUO, L., DAVIS, I., CHEN, L. & LIU, A. 2013. The power of two: arginine 51 and arginine 239* from a neighboring subunit are essential for catalysis in alpha-amino-beta-carboxymuconate-epsilon-semialdehyde decarboxylase. *J Biol Chem,* 288**,** 30862-30871.

HUO, L., LIU, F., IWAKI, H., LI, T., HASEGAWA, Y. & LIU, A. 2014. Human *α*-amino-*β*-carboxymuconate-*ε*-semialdehyde decarboxylase (ACMSD): A structural and mechanistic unveiling. *Proteins,* 83**,** 178-187.

INFORMATION HANDLING SERVICES. 2014. *Methyl Methacrylate (MMA) process summary* [Online]. Available: https://www.ihs.com/products/SRICReport-PEPReview2014-05-MethylMethacrylateMMAProcessSummary.html [Accessed March 2015].

INOUYE, M., ARNHEIM, N. & STERNGLANZ, R. 1973. Bacteriophage T7 lysozyme is an N-acetylmuramyl-L-alanine amidase. *J Biol Chem,* 248**,** 7247-7252.

INVESTMENTMINE. 2015. *5 Year Crude Oil Prices and Price Charts* [Online]. Available: http://www.infomine.com/investment/metal-prices/crude-oil/5-year/ [Accessed March].

IUBMB. Available: http://www.chem.qmul.ac.uk/iubmb/enzyme/EC4/1/1/ [Accessed September 2014].

IWAKI, H. & HASEGAWA, Y. 2007. Degradation of 2-nitrobenzoate by *Burkholderia terrae* strain KU-15. *Biosci Biotechnol Biochem,* 71**,** 145-151.

JAKLITSCH, W. M., KUBICEK, C. P. & SCRUTTON, M. C. 1991. The subcellular organization of itaconate biosynthesis in *Aspergillus terreus*. *J Gen Microbiol,* 137**,** 533-539.

JAMES, C. L. & PEI-CHING, C. 2010. *Genetically modified microorganisms for producing itaconic acid with high yields*. US2010/0285546A1.

JAMSHIDI, S., RAFII-TABAR, H. & JALILI, S. 2014. Investigation into mechanism of orotidine 5'-monophosphate decarboxylase enzyme by MM-PBSA/MM-GBSA and molecular docking. *Mol Simul,* 40**,** 469-476.

JAMSHIDI, S., JALILI, S. & RAFII-TABAR, H. 2015. Study of orotidine 5'-monophosphate decarboxylase in complex with the top three OMP, BMP, and PMP ligands by molecular dynamics simulation. *J Biomol Struct Dyn,* 33**,** 404-417.

JANES, L. E., LOWENDAHL, A. C. & KAZLAUSKAS, R. J. 1998. Quantitative screening of hydrolase libraries using pH indicators: identifying active and enantioselective hydrolases. *Chemistry,* 4**,** 2324-2331.

JANG, H.-J., YOON, S.-H., RYU, H.-K., KIM, J.-H., WANG, C.-L., KIM, J.-Y., OH, D.-K. & KIM, S.-W. 2011. Retinoid production using metabolically engineered *Escherichia coli* with a two-phase culture system. *Microb Cell Fact,* 10**,** 59.

JENSEN, R. A. 1976. Enzyme recruitment in evolution of new function. *Annu Rev Microbiol,* 30**,** 409-425.

JOCHENS, H. & BORNSCHEUER, U. T. 2010. Natural diversity to guide focused directed evolution. *ChemBioChem,* 11**,** 1861-1866.

JORDAN, F. 2004. Biochemistry: How active sites communicate in thiamine enzymes. *Science,* 306**,** 818-820.

JORDAN, F. & PATEL, H. 2013. Catalysis in enzymatic decarboxylations: Comparison of selected cofactor-dependent and cofactor-independent examples. *ACS Catal,* 3**,** 1601-1617.

KANAMASA, S., DWIARTI, L., OKABE, M. & PARK, E. Y. 2008. Cloning and functional characterization of the *cis*-aconitic acid decarboxylase (CAD) gene from *Aspergillus terreus*. *Appl Microbiol Biotechnol,* 80**,** 223-229.

KATO, Y. & ASANO, Y. 1997. 3-methylaspartate ammonia-lyase as a marker enzyme of the mesaconate pathway for (*S*)-glutamate fermentation in Enterobacteriaceae. *Arch Microbiol,* 168**,** 457-463.

KIM, J.-H., KIM, S.-W., NGUYEN, D. Q.-A., LI, H., KIM, S.-B., SEO, Y.-G., YANG, J.-K., CHUNG, I.-Y., KIM, D.-H. & KIM, C.-J. 2009. Production of β-carotene by

recombinant *Escherichia coli* with engineered whole mevalonate pathway in batch and fed-batch cultures. *Biotechnol Bioprocess Eng,* 14**,** 559-564.

KIRBY, J. & KEASLING, J. D. 2009. Biosynthesis of plant isoprenoids: perspectives for microbial engineering. *Annu Rev Plant Biol,* 60**,** 335-355.

KOBAYASHI, T. & NAKAMURA, I. 1964. Dynamics in mycelial concentration of *Aspergillus terreus* K26 in steady state of continuous culture. *J Ferm Technol,* 44**,** 264-274.

KOONTZ, W. A. & SHIMAN, R. 1976. Beef kidney 3-hydroxyanthranilic acid oxygenase. Purification, characterization, and analysis of the assay. *J Biol Chem,* 251**,** 368-377.

KOURIST, R., GUTERL, J.-K., MIYAMOTO, K. & SIEBER, V. 2014. Enzymatic decarboxylation-an emerging reaction for chemicals production from renewable resources. *ChemCatChem,* 6**,** 689-701.

KREPKIY, D. V. & MIZIORKO, H. M. 2004. Identification of active site residues in mevalonate diphosphate decarboxylase: Implications for a family of phosphotransferases. *Protein Sci,* 13**,** 1875-1881.

KREPKIY, D. V. & MIZIORKO, H. M. 2005. Investigation of the functional contributions of invariant Serine residues in yeast mevalonate diphosphate decarboxylase. *Biochemistry,* 44**,** 2671-2677.

KUCHNER, O. & ARNOLD, F. H. 1997. Directed evolution of enzyme catalysts. *Trends Biotechnol,* 15**,** 523-530.

KUIPERS, R. K., JOOSTEN, H.-J., VAN BERKEL, W. J. H., LEFERINK, N. G. H., ROOIJEN, E., ITTMANN, E., VAN ZIMMEREN, F., JOCHENS, H., BORNSCHEUER, U., VRIEND, G., MARTINS DOS SANTOS, V. A. P. & SCHAAP, P. J. 2010. 3DM: Systematic analysis of heterogeneous superfamily data to discover protein functionalities. *Proteins,* 78**,** 2101-2113.

KURIOKA, S. & MATSUDA, M. 1976. Phospholipase C assay using p-nitrophenylphosphorylcholine together with sorbitol and its application to studying the metal and detergent requirement of the enzyme. *Anal Biochem,* 75**,** 281-289.

KUTNER, A., CHODYNSKI, M. & HALKES, S. J. 1996. Citramalic acid as chiral synthon, synthesis on chiral aliphatic sulfones. *Synth Commun,* 26**,** 1175-1182.

KUZUYAMA, T. & SETO, H. 2012. Two distinct pathways for essential metabolic precursors for isoprenoid biosynthesis. *Proc Jpn Acad Ser B Phys Biol Sci,* 88**,** 41-52.

LEE, J. K. 1997. A Proficient Enzyme Revisited: The Predicted Mechanism for Orotidine Monophosphate Decarboxylase. *Science,* 276**,** 942-945.

LEFURGY, S. T., RODRIGUEZ, S. B., PARK, C. S., CAHILL, S., SILVERMAN, R. B. & LEYH, T. S. 2010. Probing ligand-binding pockets of the mevalonate pathway enzymes from *Streptococcus pneumoniae*. *J Biol Chem,* 285**,** 20654-20663.

LERCH, T. Z., COUCHENEY, E. & HERRMANN, A. M. 2013. Sensitivity of soil microbial catabolic profiles to a gradient of carbon inputs: Does the soil organic matter matter? *Soil Biol Biochem,* 57**,** 911-915.

LEWIN, R., GOODALL, M., THOMPSON, M. L., LEIGH, J., BREUER, M., BALDENIUS, K. & MICKLEFIELD, J. 2015. Enzymatic Enantioselective Decarboxylative Protonation of Heteroaryl Malonates. *Chem - Eur J,* 21**,** 6557-6563.

LEYS, D., PAYNE, K. A. P., SCRUTTON, N. S., PARKER, D. A. & MURPHY, A. J. 2013. *Saccharomyces cerevisiae ferulic acid decarboxylase and Ph acrylic acid decarboxylase, their sequences, recombinant production and use in preparing mono-unsaturated alkene or in generation of biofuel*. US20130330795A1.

LI, A., VAN LUIJK, N., TER BEEK, M., CASPERS, M., PUNT, P. & VAN DER WERF, M. 2011. A clone-based transcriptomics approach for the identification of genes relevant for itaconic acid production in *Aspergillus*. *Fungal Genet Biol,* 48**,** 602-611.

LI, A., PFELZER, N., ZUIJDERWIJK, R. & PUNT, P. 2012a. Enhanced itaconic acid production in *Aspergillus niger* using genetic modification and medium optimization. *BMC Biotechnol,* 12**,** 57.

LI, A., CASPERS, M. & PUNT, P. 2013a. A systems biology approach for the identification of target genes for the improvement of itaconic acid production in *Aspergillus species*. *BMC Res Notes,* 6**,** 505.

LI, A., PFELZER, N., ZUIJDERWIJK, R., BRICKWEDDE, A., VAN ZEIJL, C. & PUNT, P. 2013b. Reduced by-product formation and modified oxygen availability improve itaconic acid production in *Aspergillus niger*. *Appl Microbiol Biotechnol,* 97**,** 3901-3911.

LI, T., WALKER, A. L., IWAKI, H., HASEGAWA, Y. & LIU, A. 2005. Kinetic and spectroscopic characterization of ACMSD from *Pseudomonas fluorescens* reveals a pentacoordinate mononuclear metallocofactor. *J Am Chem Soc,* 127**,** 12282-12290.

LI, T., IWAKI, H., FU, R., HASEGAWA, Y., ZHANG, H. & LIU, A. 2006. α-amino-β-carboxymuconic-ε-semialdehyde decarboxylase (ACMSD) is a new member of the amidohydrolase superfamily. *Biochemistry,* 45**,** 6628-6634.

LI, T., MA, J. K., HOSLER, J. P., DAVIDSON, V. L. & LIU, A. 2007. Detection of transient intermediates in the metal-dependent nonoxidative decarboxylation catalyzed by α-amino-β-carboxymuconate-ε-semialdehyde decarboxylase. *J Am Chem Soc,* 129**,** 9278-9279.

LI, T., HUO, L., PULLEY, C. & LIU, A. 2012b. Decarboxylation mechanisms in biological system. *Bioorg Chem,* 43**,** 2-14.

LIE, M. A., CELIK, L., JØRGENSEN, K. A. & SCHIØTT, B. 2005. Cofactor Activation and Substrate Binding in Pyruvate Decarboxylase. Insights into the Reaction Mechanism from Molecular Dynamics Simulations. *Biochemistry,* 44**,** 14792-14806.

LIESE, A., SEELBACH, K., BUCHHOLZ, A., HABERLAND, J., LIESE, A., SEELBACH, K. & WANDREY, C. 2007. Processes. *Industrial Biotransformations.* Wiley-VCH Verlag GmbH.

LIN, F., FERGUSON, K. L., BOYER, D. R., LIN, X. N. & MARSH, E. N. G. 2015. Isofunctional enzymes PAD1 and UbiX catalyze formation of a novel cofactor required by ferulic acid decarboxylase and 4-hydroxy-3-polyprenylbenzoic acid decarboxylase. *ACS Chem Biol,* 10**,** 1137-1144.

LIND, M. & HIMO, F. 2014. Theoretical Study of Reaction Mechanism and Stereoselectivity of Arylmalonate Decarboxylase. *ACS Catal,* 4**,** 4153-4160.

LINGEN, B., KOLTER-JUNG, D., DÜNKELMANN, P., FELDMANN, R., GRÖTZINGER, J., POHL, M. & MÜLLER, M. 2003. Alteration of the substrate specificity of benzoylformate decarboxylase from *Pseudomonas putida* by directed evolution. *ChemBioChem,* 4**,** 721-726.

LIU, A. & ZHANG, H. 2006. Transition metal-catalyzed nonoxidative decarboxylation reactions. *Biochemistry,* 45**,** 10407-10411.

LIU, D., KARSTEN, W. E. & COOK, P. F. 2000. Lysine 199 is the general acid in the NAD-malic enzyme reaction. *Biochemistry,* 39**,** 11955-11960.

LOHKAMP, B., BÄUERLE, B., RIEGER, P.-G. & SCHNEIDER, G. 2006. Three-dimensional structure of iminodisuccinate epimerase defines the fold of the MmgE/PrpD protein family. *J Mol Biol,* 362**,** 555-566.

LUCITE INTERNATIONAL. 2012. *Taking a closer look at the MMA markets* [Online]. Available: http://www.luciteinternational.com/taking-a-closer-look-at-the-mma-markets-11/ [Accessed October 2014].

LUCITE INTERNATIONAL. 2014. *Methacrylate monomers -Innovation* [Online]. Available: http://www.luciteinternational.com/monomers-eame-responsibility-innovation-33/ [Accessed March 2015].

LUTZ, S. 2010. Beyond directed evolution—semi-rational protein engineering and design. *Curr Opin Biotechnol,* 21**,** 734-743.

MARLIERE, P. 2010. *Production of alkenes by enzymatic decarboxylation of 3-hydroxyalkanoic acids*. WO2010001078A2.

MARLIERE, P. 2011. *Enymatic production of 3-hydroxy-3-methylbutyric acid from acetone and acetyl-Coa*. WO2011032934A1.

MARLIERE, P., ANISSIMOVA, M., CHAYOT, R. & DELCOURT, M. 2011. *Process for the production of isoprenol from mevalonate employing a diphosphomevalonate decarboxylase*. WO2011076261A1.

MARLIERE, P., DELCOURT, M., ANISSIMOVA, M. & TALLON, R. 2012. *Production of alkenes by combined enzymatic conversion of 3-hydroxyalkanoic acids using different mevalonate pyrophosphate decarboxylases*. WO2012052427A1.

MARLIERE, P., ANISSIMOVA, M. & ALLARD, M. 2014. *Production of alkenes from 3-hydroxy-1-carboxylic acids via 3-sulfonyloxy-1-carboxylic acids*. WO2014064198A1.

MARRS, B., DELAGRAVE, S. & MURPHY, D. 1999. Novel approaches for discovering industrial enzymes. *Curr Opin Microbiol,* 2**,** 241-245.

MARSHALL, V. D. & SOKATCH, J. R. 1972. Regulation of valine catabolism in *Pseudomonas putida*. *J Bacteriol,* 110**,** 1073-1081.

MARTYNOWSKI, D., EYOBO, Y., LI, T., YANG, K., LIU, A. & ZHANG, H. 2006. Crystal structure of α-amino-β-carboxymuconate-ε-semialdehyde decarboxylase: Insight into the active site and catalytic mechanism of a novel decarboxylation reaction. *Biochemistry,* 45**,** 10412-10421.

MASSEY, L. K., SOKATCH, J. R. & CONRAD, R. S. 1976. Branched-chain amino acid catabolism in bacteria. *Bacteriol Rev,* 40**,** 42-54.

MATA, L., GRIPON, J.-C. & MISTOU, M.-Y. 1999. Deletion of the four C-terminal residues of PepC converts an aminopeptidase into an oligopeptidase. *Protein Eng,* 12**,** 681-686.

MCCARTHY, J. G., EISMAN, E. B., KULKARNI, S., GERWICK, L., GERWICK, W. H., WIPF, P., SHERMAN, D. H. & SMITH, J. L. 2012. Structural basis of functional group activation by sulfotransferases in complex metabolic pathways. *ACS Chem Biol,* 7**,** 1994-2003.

MCCULLUM, E. O., WILLIAMS, B. A., ZHANG, J. & CHAPUT, J. C. 2010. Random mutagenesis by error-prone PCR. *Methods Mol Biol,* 634**,** 103-109.

MCKENNA, R. & NIELSEN, D. R. 2011. Styrene biosynthesis from glucose by engineered *E. coli*. *Metab Eng,* 13**,** 544-554.

MCKENNA, R., THOMPSON, B., PUGH, S. & NIELSEN, D. R. 2014. Rational and combinatorial approaches to engineering styrene production by *Saccharomyces cerevisiae*. *Microb Cell Fact,* 13**,** 123.

MEDICI, R., DE MARIA, P. D., OTTEN, L. G. & STRAATHOF, A. J. J. 2011. *A High-Throughput Screening Assay for Amino Acid Decarboxylase Activity*.

MEHLER, A. H., YANO, K. & MAY, E. L. 1964. Nicotinic acid biosynthesis. Control by an enzyme that competes with a spontaneous reaction. *Science* 145**,** 817-819.

MENDEZ-PEREZ, D., BEGEMANN, M. B. & PFLEGER, B. F. 2011. Modular synthase-encoding gene involved in α-olefin biosynthesis in *Synechococcus sp.* strain PCC 7002. *Appl Environ Microbiol,* 77**,** 4264-4267.

METACYC. 2015. *EC 2.7.1 -- Phosphotransferases with an alcohol group as acceptor* [Online]. Available: http://metacyc.org/META/NEW-IMAGE?type=EC-NUMBER&object=EC-2.7.1 [Accessed March 2015].

METHACRYLATE PRODUCERS ASSOCIATION. 2012. *About methacrylates* [Online]. Available: http://www.mpausa.org/about-methacrylates2/ [Accessed February 2015].

MICHIHARA, A., AKASAKI, K., YAMORI, Y. & TSUJI, H. 2002. Purification and characterization of mouse mevalonate pyrophosphate decarboxylase. *Biol Pharm Bull,* 25**,** 302-306.

MISAWA, N. 2011. Pathway engineering for functional isoprenoids. *Curr Opin Biotechnol,* 22**,** 627-633.

MITSUBISHI RAYON CO. 2011. *Development of production technology for sustainable MMA* [Online]. Available: https://www.mrc.co.jp/english/pressroom/p11/press20111109.html [Accessed March 2015].

MIYAMOTO, K. & OHTA, H. 1992. Purification and properties of a novel arylmalonate decarboxylase from *Alcaligenes bronchisepticus* KU 1201. *Eur J Biochem,* 210**,** 475-481.

MIYAMOTO, K. 2014. Altered function of arylmalonate decarboxylase decarboxylase: dramatic improvement in activity and reversal of stereoselectivity due to resonable design. *Seibutsu Kogaku Kaishi,* 92**,** 424-427.

MIYAUCHI, Y., KOURIST, R., UEMURA, D. & MIYAMOTO, K. 2011. Dramatically improved catalytic activity of an artificial (*S*)-selective arylmalonate decarboxylase by structure-guided directed evolution. *Chemical Communications,* 47**,** 7503.

MOFFATT, B. A. & STUDIER, F. W. 1987. T7 lysozyme inhibits transcription by T7 RNA polymerase. *Cell,* 49**,** 221-227.

MORLEY, K. L. & KAZLAUSKAS, R. J. 2005. Improving enzyme properties: when are closer mutations better? *Trends Biotechnol,* 23**,** 231-237.

MORLEY, K. L., GROSSE, S., LEISCH, H. & LAU, P. C. K. 2013. Antioxidant canolol production from a renewable feedstock via an engineered decarboxylase. *Green Chem.,* 15**,** 3312-3317.

MUKAI, N., MASAKI, K., FUJII, T., KAWAMUKAI, M. & IEFUJI, H. 2010. PAD1 and FDC1 are essential for the decarboxylation of phenylacrylic acids in *Saccharomyces cerevisiae*. *J Biosci Bioeng,* 109**,** 564-569.

MURAKI, T., TAKI, M., HASEGAWA, Y., IWAKI, H. & LAU, P. C. K. 2003. Prokaryotic homologs of the eukaryotic 3-hydroxyanthranilate 3,4-dioxygenase and 2-amino-3-carboxymuconate-6-semialdehyde decarboxylase in the 2-nitrobenzoate degradation pathway of *Pseudomonas fluorescens* strain KU-7. *Appl Environ Microbiol,* 69**,** 1564-1572.

NAGAI, K. 2001. New developments in the production of methyl methacrylate. *Applied Catalysis A: General,* 221**,** 367-377.

NELSON, D. L., LEHNINGER, A. L. & COX, M. M. 2008. *Lehninger principles of biochemistry (5th ed.),* New York, W.H. Freeman.

NEYLON, C. 2004. Chemical and biochemical strategies for the randomization of protein encoding DNA sequences: library construction methods for directed evolution. *Nucleic Acids Res,* 32**,** 1448-1459.

NISHIZUKA, Y., ICHIYAMA, A. & HAYAISHI, O. 1970. Metabolism of the benzene ring of tryptophan (mammals). *In:* HERBERT TABOR, C. W. T. (ed.) *Methods in Enzymology.* Academic Press.

NOBILI, A., GALL, M. G., PAVLIDIS, I. V., THOMPSON, M. L., SCHMIDT, M. & BORNSCHEUER, U. T. 2013. Use of "small but smart" libraries to enhance the enantioselectivity of an esterase from *Bacillus stearothermophilus* towards tetrahydrofuran-3-yl acetate. *FEBS J,* 280**,** 3084-3093.

NOVAGEN 2003. *pET system Manual,* Darmstadt.

OBENAUS, F., DROSTE, W. & NEUMEISTER, J. 2000. Butenes. *Ullmann's Encyclopedia of Industrial Chemistry.* Wiley-VCH Verlag GmbH & Co. KGaA.

OECD. 2003. *SIDS initial assessment report: isobutylene* [Online]. Available: http://www.inchem.org/documents/sids/sids/115117.pdf [Accessed March 2015].

OKABE, M., LIES, D., KANAMASA, S. & PARK, E. Y. 2009. Biotechnological production of itaconic acid and its biosynthesis in *Aspergillus terreus*. *Appl Microbiol Biotechnol,* 84**,** 597-606.

OKAMOTO, S., CHIN, T., HIRATSUKA, K., ASO, Y., TANAKA, Y., TAKAHASHI, T. & OHARA, H. 2014. Production of itaconic acid using metabolically engineered *Escherichia coli*. *J Gen Appl Microbiol,* 60**,** 191-197.

OKRASA, K., LEVY, C., WILDING, M., GOODALL, M., BAUDENDISTEL, N., HAUER, B., LEYS, D. & MICKLEFIELD, J. 2009. Structure-Guided Directed Evolution of Alkenyl and Arylmalonate Decarboxylases. *Angewandte Chemie International Edition,* 48**,** 7691-7694.

OSTERHOUT, R. E., BURGARD, A. P. & BURK, M. J. 2013. *Metabolically engineered microorganisms and pathways for fermentative prodn. of 2,4-pentadienoate, butadiene, propylene, 1,3-butanediol and related alcohols*. WO2013028519A1.

OTTE, K. B. & HAUER, B. 2015. Enzyme engineering in the context of novel pathways and products. *Curr Opin Biotechnol,* 35**,** 16-22.

OTTEN, A., BROCKER, M. & BOTT, M. 2015. Metabolic engineering of corynebacterium glutamicum for the production of itaconate. *Metab Eng*.

PANG, Y., SHEN, G.-A., BERGES, T., CARDIER, H., WU, W., SUN, X. & TANG, K. 2006. Molecular cloning, characterization and heterologous expression in *Saccharomyces cerevisiae* of a mevalonate diphosphate decarboxylase cDNA from *Ginkgo biloba*. *Physiol Plant,* 127**,** 19-27.

PAPWORTH, C., BAUER, J. C., BRAMAN, J. & WRIGHT, D. A. 1996. Site-directed mutagenesis in one day with >80% efficiency. *Strategies,* 9**,** 3-4.

PAYNE, K. A., WHITE, M. D., FISHER, K., KHARA, B., BAILEY, S. S., PARKER, D., RATTRAY, N. J., TRIVEDI, D. K., GOODACRE, R., BEVERIDGE, R., BARRAN, P., RIGBY, S. E., SCRUTTON, N. S., HAY, S. & LEYS, D. 2015. New cofactor supports alpha,beta-unsaturated acid decarboxylation via 1,3-dipolar cycloaddition. *Nature,* 522**,** 497-501.

PIRES, R. & BURGER, K. 1996. A new method for the synthesis of (5')n-methylisoserine and its incorporation into peptides. *Synthesis,* 11**,** 1277-1280.

PLUMRIDGE, A., STRATFORD, M., LOWE, K. C. & ARCHER, D. B. 2007. The weak-acid preservative sorbic acid is decarboxylated and detoxified by a phenylacrylic acid decarboxylase, PadA1, in the spoilage mold *Aspergillus niger*. *Appl Environ Microbiol,* 74**,** 550-552.

PLUMRIDGE, A., MELIN, P., STRATFORD, M., NOVODVORSKA, M., SHUNBURNE, L., DYER, P. S., ROUBOS, J. A., MENKE, H., STARK, J., STAM, H. & ARCHER, D. B. 2010. The decarboxylation of the weak-acid preservative, sorbic acid, is encoded by linked genes in *Aspergillus spp*. *Fungal Genet Biol,* 47**,** 683-692.

PRASAD, S., BOCOLA, M. & REETZ, M. T. 2011. Revisiting the lipase from *Pseudomonas aeruginosa*: directed evolution of substrate acceptance and enantioselectivity using iterative saturation mutagenesis. *Chemphyschem,* 12**,** 1550-1557.

PRESSMAN, D. & LUCAS, H. J. 1940. The hydration of unsaturated compounds. VIII. The rate of hydration of β,β-dimethylacrylic acid; the rates of dehydration and decarboxylation of β-hydroxyisovaleric acid. *J Am Chem Soc,* 62**,** 2069-2080.

PUCCI, L., PEROZZI, S., CIMADAMORE, F., ORSOMANDO, G. & RAFFAELLI, N. 2007. Tissue expression and biochemical characterization of human 2-amino 3-carboxymuconate 6-semialdehyde decarboxylase, a key enzyme in tryptophan catabolism. *FEBS J,* 274**,** 827-840.

PYO, S.-H., DISHISHA, T., DAYANKAC, S., GERELSAIKHAN, J., LUNDMARK, S., REHNBERG, N. & HATTI-KAUL, R. 2012. A new route for the synthesis of methacrylic acid from 2-methyl-1,3-propanediol by integrating biotransformation and catalytic dehydration. *Green Chem,* 14**,** 1942-1948.

QIU, Y. & LI, D. 2006. Inhibition of mevalonate 5-diphosphate decarboxylase by fluorinated substrate analogs. *Biochim Biophys Acta,* 1760**,** 1080-1087.

QIU, Y., GAO, J., GUO, F., QIAO, Y. & LI, D. 2007. Mutation and inhibition studies of mevalonate 5-diphosphate decarboxylase. *Bioorg Med Chem Lett,* 17**,** 6164-6168.

RAJENDRAPRASAD, N. & BASAVAIAH, K. 2009. Determination of olanzapine by spectrophotometry using permanganate. *Braz J Pharm Sci,* 45**,** 539-550.

RALSER, M., KUHL, H., RALSER, M., WERBER, M., LEHRACH, H., BREITENBACH, M. & TIMMERMANN, B. 2012. The *Saccharomyces cerevisiae* W303-K6001 cross-platform genome sequence: insights into ancestry and physiology of a laboratory mutt. *Open Biol,* 2**,** 120093.

RANGARAJAN, E. S., LI, Y., IANNUZZI, P., TOCILJ, A., HUNG, L. W., MATTE, A. & CYGLER, M. 2008. Crystal structure of a dodecameric FMN-dependent UbiX-like decarboxylase (Pad1) from *Escherichia coli* O157: H7. *Protein Sci,* 13**,** 3006-3016.

RASILA, T. S., PAJUNEN, M. I. & SAVILAHTI, H. 2009. Critical evaluation of random mutagenesis by error-prone polymerase chain reaction protocols, *Escherichia coli* mutator strain, and hydroxylamine treatment. *Anal Biochem,* 388**,** 71-80.

REARDON, J. E. A., ROBERT H. 1987. Inhibition of cholesterol biosynthesis by fluorinated mevalonate analogs. *Biochemistry,* 26**,** 4717-4722.

REETZ, M. T., BOCOLA, M., CARBALLEIRA, J. D., ZHA, D. & VOGEL, A. 2005. Expanding the range of substrate acceptance of enzymes: combinatorial active-site saturation test. *Angew Chem Int Ed Engl,* 44**,** 4192-4196.

REETZ, M. T., CARBALLEIRA, J. D. & VOGEL, A. 2006a. Iterative saturation mutagenesis on the basis of B factors as a strategy for increasing protein thermostability. *Angew Chem Int Ed Engl,* 45**,** 7745-7751.

REETZ, M. T., PEYRALANS, J. J., MAICHELE, A., FU, Y. & MAYWALD, M. 2006b. Directed evolution of hybrid enzymes: Evolving enantioselectivity of an achiral Rh-complex anchored to a protein. *Chem Commun (Camb)***,** 4318-4320.

REETZ, M. T., WANG, L. W. & BOCOLA, M. 2006c. Directed evolution of enantioselective enzymes: Iterative cycles of CASTing for probing protein-sequence space. *Angew Chem Int Ed Engl,* 45**,** 1236-1241.

REETZ, M. T. & CARBALLEIRA, J. D. 2007. Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes. *Nat Protoc,* 2**,** 891-903.

REETZ, M. T., KAHAKEAW, D. & LOHMER, R. 2008. Addressing the numbers problem in directed evolution. *Chembiochem,* 9**,** 1797-1804.

REETZ, M. T. & SANCHIS, J. 2008. Constructing and analyzing the fitness landscape of an experimental evolutionary process. *Chembiochem,* 9**,** 2260-2267.

REETZ, M. T., PRASAD, S., CARBALLEIRA, J. D., GUMULYA, Y. & BOCOLA, M. 2010a. Iterative saturation mutagenesis accelerates laboratory evolution of enzyme stereoselectivity: Rigorous comparison with traditional methods. *J Am Chem Soc,* 132**,** 9144-9152.

REETZ, M. T., SONI, P., FERNANDEZ, L., GUMULYA, Y. & CARBALLEIRA, J. D. 2010b. Increasing the stability of an enzyme toward hostile organic solvents by directed evolution based on iterative saturation mutagenesis using the B-FIT method. *Chem Commun (Camb),* 46**,** 8657-8658.

REETZ, M. T. 2011. Laboratory evolution of stereoselective enzymes: A prolific source of catalysts for asymmetric reactions. *Angew Chem Int Ed Eng,* 50**,** 138-174.

REETZ, M. T. 2013. Biocatalysis in organic chemistry and biotechnology: past, present, and future. *J Am Chem Soc,* 135**,** 12480-12496.

REINHARD, J. F., JR. 2004. Pharmacological manipulation of brain kynurenine metabolism. *Ann N Y Acad Sci,* 1035**,** 335-349.

RENAULT, P., BEN-SASSI, M. & BÉRARD, A. 2013. Improving the MicroResp™ substrate-induced respiration method by a more complete description of CO2 behavior in closed incubation wells. *Geoderma,* 207-208**,** 82-91.

REUTHER, G., HARRIS, R., GIRVIN, M. & LEYH, T. S. 2010. Backbone 1H, 13C, 15N NMR assignments of the unliganded and substrate ternary complex forms of mevalonate diphosphate decarboxylase from *Streptococcus pneumoniae*. *Biomol NMR Assign,* 5**,** 11-14.

REYMOND, J.-L. 2005. *Enzyme assays: high-throughput screening, genetic selection and fingerprinting*, Wiley-VCH, Weinheim.

ROBERTS, S. M., TURNER, N. J., WILLETTS, A. J. & TURNER, M. K. 1995. *Introduction to Biocatalysis Using Enzymes and Microorganisms*, Cambridge University Press.

ROBINSON, W. G., NAGLE, R., BACHHAWAT, B. K., KUPIECKI, F. P. & COON, M. J. 1957. Coenzyme A thiol esters of isobutyric, methacrylic, and *β*-hydroxyisobutyric acid as intermediates in the enzymati degradation of valine *J Biol Chem,* 224**,** 1-11.

RODRÍGUEZ, H., ANGULO, I., DE LAS RIVAS, B., CAMPILLO, N., PÁEZ, J. A., MUÑOZ, R. & MANCHEÑO, J. M. 2010. p-Coumaric acid decarboxylase from *Lactobacillus plantarum*: Structural insights into the active site and decarboxylation catalytic mechanism. *Proteins,* 78**,** 1662-1676.

ROHMER, M., GROSDEMANGE-BILLIARD, C., SEEMANN, M. & TRITSCH, D. 2004. Isoprenoid biosynthesis as a novel target for antibacterial and antiparasitic drugs. *Cur Opin Investig Drugs,* 5**,** 154-162.

ROSENBERG, R. M., HERREID, R. M., PIAZZA, G. J. & O'LEARY, M. H. 1989. Indicator assay for amino acid decarboxylases. *Anal Biochem,* 181**,** 59-65.

ROY, A., KUCUKURAL, A. & ZHANG, Y. 2010. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc,* 5**,** 725-738.

RUFF, A. J., DENNIG, A. & SCHWANEBERG, U. 2013. To get what we aim for - progress in diversity generation methods. *FEBS J,* 280**,** 2961-2978.

RUSSELL, A. J. & FERSHT, A. R. 1987. Rational modification of enzyme catalysis by engineering surface charge. *Nature,* 328**,** 496-500.

SAKAMURI, R. 2000. Esters, Organic. *Kirk-Othmer Encyclopedia of Chemical Technology.* John Wiley & Sons, Inc.

SANGER, F. & COULSON, A. R. 1975. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J Mol Biol,* 94**,** 441-448.

SATO, E., YAMAZAKI, M., NAKAJIMA, E., YU, F., FUJITA, T. & MIZUNASHI, W. 2014a. *Method for manufacturing methacrylic acid and/or ester thereof with microorganisms or enzymes (from biomass)*. WO2014038216A1.

SATO, E., YU, F., MIZUNASHI, W. & NAKAJIMA, E. 2014b. *Enzymic preparation of producing methacrylic acid ester*. WO2014038214A1.

SCALCINATI, G., KNUF, C., PARTOW, S., CHEN, Y., MAURY, J., SCHALK, M., DAVIET, L., NIELSEN, J. & SIEWERS, V. 2012. Dynamic control of gene expression in *Saccharomyces cerevisiae* engineered for the production of plant sesquiterpene α-santalene in a fed-batch mode. *Metab Eng,* 14**,** 91-103.

SCHWARCZ, R. 2004. The kynurenine pathway of tryptophan degradation as a drug target. *Curr Opin Pharmacol,* 4**,** 12-17.

SHI, L., QIN, L., XU, Y., REN, A., FANG, X., MU, D., TAN, Q. & ZHAO, M. 2011. Molecular cloning, characterization, and function analysis of a mevalonate pyrophosphate decarboxylase gene from *Ganoderma lucidum*. *Mol Biol Rep,* 39**,** 6149-6159.

SHIBATA, K., FUKUWATARI, T. & SASAKI, R. 2007. Phthalate esters enhance quinolinate production by inhibiting amino-carboxymuconate-semialdehyde decarboxylase (ACMSD), a key enzyme of the tryptophan-niacin pathway. *Int Congr Ser,* 1304**,** 184-194.

SMOLDERS, K. & BAEYENS, J. 2004. Thermal degradation of PMMA in fluidised beds. *Waste management,* 24**,** 849-857.

SONNTAG, F., BUCHHAUPT, M. & SCHRADER, J. 2014. Thioesterases for ethylmalonyl–CoA pathway derived dicarboxylic acid production in *Methylobacterium extorquens* AM1. *Appl Microbiol Biotechnol,* 98**,** 4533-4544.

STEIGER, M. G., BLUMHOFF, M. L., MATTANOVICH, D. & SAUER, M. 2013. Biochemistry of microbial itaconic acid production. *Front Microbiol,* 4**,** 23.

STEMMER, W. P. 1994. DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. *Proc Natl Acad Sci* 91**,** 10747–10751.

STRATFORD, M., PLUMRIDGE, A. & ARCHER, D. B. 2007. Decarboxylation of sorbic acid by spoilage yeasts is associated with the PAD1 Gene. *Appl Environ Microbiol,* 73**,** 6534-6542.

STRATFORD, M., PLUMRIDGE, A., PLEASANTS, M. W., NOVODVORSKA, M., BAKER-GLENN, C. A. G., PATTENDEN, G. & ARCHER, D. B. 2012. Mapping the structural requirements of inducers and substrates for decarboxylation of weak acid preservatives by the food spoilage mould *Aspergillus niger*. *Int J Food Microbiol,* 157**,** 375-383.

STRAUSS, E., ZHAI, H., BRAND, L. A., MCLAFFERTY, F. W. & BEGLEY, T. P. 2004. Mechanistic studies on phosphopantothenoylcysteine decarboxylase: Trapping of an enethiolate intermediate with a mechanism-based inactivating agent. *Biochemistry,* 43**,** 15520-15533.

STUDIER, F. W. 1991. Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. *J Mol Biol,* 219**,** 37-44.

SWALLOW, M. J. B. & QUIDEAU, S. A. 2015. A method for determining community level physiological profiles of organic soil horizons. *Soil Sci. Soc. Am. J.,* 79**,** 536-542.

TAMURA, K., STECHER, G., PETERSON, D., FILIPSKI, A. & KUMAR, S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol,* 30**,** 2725-2729.

TANABE, A., EGASHIRA, Y., FUKUOKA, S.-I., SHIBATA, K. & SANADA, H. 2002. Purification and molecular cloning of rat 2-amino-3-carboxymuconate-6-semialdehyde decarboxylase. *Biochem J,* 361**,** 567-575.

TANNER, A., BOWATER, L., FAIRHURST, S. A. & BORNEMANN, S. 2001. Oxalate decarboxylase requires manganese and dioxygen for activity: Overexpression and characterization of *Bacillus subtilis* YvrK and YoaN. *J Biol Chem,* 276**,** 43627-43634.

TAO, X., YANG, Z. & TONG, L. 2003. Crystal structures of substrate complexes of malic enzyme and insights into the catalytic mechanism. *Structure,* 11**,** 1141-1150.

TEVŽ, G., BENČINA, M. & LEGIŠA, M. 2010. Enhancing itaconic acid production by *Aspergillus terreus*. *Appl Microbiol Biotechnol,* 87**,** 1657-1664.

TONEY, M. D. 2005. Reaction specificity in pyridoxal phosphate enzymes. *Arch Biochem Biophys,* 433**,** 279-287.

TOOGOOD, H. S. & SCRUTTON, N. S. 2013. Enzyme engineering toolbox – a 'catalyst' for change. *Catal Sci Technol,* 3**,** 2182.

TROTT, O. & OLSON, A. J. 2010. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem,* 31**,** 455-461.

TURNER, N. J. 2009. Directed evolution drives the next generation of biocatalysts. *Nature Chemical Biology,* 5**,** 567-573.

TURNER, N. J. & TRUPPO, M. D. 2013. Biocatalysis enters a new era. *Curr Opin Chem Biol,* 17**,** 212-214.

UEDA, M., SASHIDA, R., YAMADA, H. & ASANO, Y. 1993. Production of (*S*)-(+)-citramalic acid from itaconic acid by resting cells of *Alcaligenes denitrificans* strain MCI2775. *Appl Microbiol Biotechnol,* 40**,** 466-469.

ULMER, K. M. 1983. Protein engineering. *Science* 219**,** 666-671.

VAN DER STRAAT, L., VERNOOIJ, M., LAMMERS, M., VAN, D. B. W., SCHONEWILLE, T., CORDEWENER, J., VAN, D. M. I., KOOPS, A. & DE, G. L. H. 2014. Expression of the *Aspergillus terreus* itaconic acid biosynthesis cluster in *Aspergillus niger*. *Microb Cell Fact,* 13**,** doi:10.1186/1475-2859-13-11.

VANLEEUWEN, B. N. M., WULP, A. M., DUIJNSTEE, I., MARIS, A. J. A. & STRAATHOF, A. J. J. 2012. Fermentative production of isobutene. *Appl Microbiol Biotechnol,* 93**,** 1377-1387.

VANNICE, J. C., SKAFF, D. A., KEIGHTLEY, A., ADDO, J. K., WYCKOFF, G. J. & MIZIORKO, H. M. 2014. Identification in *Haloferax volcanii* of phosphomevalonate decarboxylase and isopentenyl phosphate kinase as catalysts of the terminal enzyme reactions in an archaeal alternate mevalonate pathway. *J Bacteriol,* 196**,** 1055-1063.

VERSÉES, W., SPAEPEN, S., VANDERLEYDEN, J. & STEYAERT, J. 2007. The crystal structure of phenylpyruvate decarboxylase from *Azospirillum brasilense* at 1.5 Å resolution: Implications for its catalytic and regulatory mechanism. *FEBS J,* 274**,** 2363-2375.

VINOKUR, J. M., KORMAN, T. P., CAO, Z. & BOWIE, J. U. 2014a. Evidence of a novel mevalonate pathway in archaea. *Biochemistry,* 53**,** 4161-4168.

VINOKUR, J. M., KORMAN, T. P., SAWAYA, M. R., COLLAZO, M., CASCIO, D. & BOWIE, J. U. 2014b. Structural analysis of mevalonate-3-kinase provides insight into the mechanisms of isoprenoid pathway decarboxylases. *Protein Sci,* 24**,** 212-220.

VINOKUR, J. M., KORMAN, T. P., SAWAYA, M. R., COLLAZO, M., CASCIO, D. & BOWIE, J. U. 2015. Structural analysis of mevalonate-3-kinase provides insight into the mechanisms of isoprenoid pathway decarboxylases. *Protein Sci,* 24**,** 212-220.

VOGELS, G. D. & VAN DER DRIFT, C. 1976. Degradation of purines and pyrimidines by microorganisms. *Bacteriol Rev,* 40**,** 403-468.

VOYNOVA, N., FU, Z., BATTAILE, K., HERDENDORF, T., KIM, J. & MIZIORKO, H. 2008. Human mevalonate diphosphate decarboxylase: characterization, investigation of the mevalonate diphosphate binding site, and crystal structure. *Arch Biochem Biophys,* 480**,** 58-67.

VUORISTO, K. S., MARS, A. E., SANGRA, J. V., SPRINGER, J., EGGINK, G., SANDERS, J. P. M. & WEUSTHUIS, R. A. 2015. Metabolic engineering of itaconate production in *Escherichia coli*. *Appl Microbiol Biotechnol,* 99**,** 221-228.

WACHSMAN, J. T. 1956. The role of α-ketoglutarate and mesaconate in glutamate fermentation by *Clostridium tetanomorphum*. *J Biol Chem,* 223**,** 19-27.

WAKELIN, S., LOMBI, E., DONNER, E., MACDONALD, L., BLACK, A. & O'CALLAGHAN, M. 2013. Application of MicroResp for soil ecotoxicology. *Environ Pollut,* 179**,** 177-184.

WANG, C., YOON, S.-H., JANG, H.-J., CHUNG, Y.-R., KIM, J.-Y., CHOI, E.-S. & KIM, S.-W. 2011. Metabolic engineering of *Escherichia coli* for α-farnesene production. *Metab Eng,* 13**,** 648-655.

WANG, J. & ZHANG, K. 2015. Production of mesaconate in *Escherichia coli* by engineered glutamate mutase pathway. *Metab Eng,* 9**,** 190-196.

WEBB, M. E. & ABELL, C. Molecular mechanism of *E. coli* aspartate decarboxylase activation is revealed by site-directed mutagenesis and crystallography. 2003. American Chemical Society, BIOL-095.

WEERASINGHE, S. & DASSANAYAKE, R. S. 2009. Simulation of structural and functional properties of mevalonate diphosphate decarboxylase (MVD). *J Mol Model,* 16**,** 489-498.

WEN, M., BOND-WATTS, B. B. & CHANG, M. C. Y. 2013. Production of advanced biofuels in engineered E. coli. *Current Opinion in Chemical Biology,* 17**,** 472-479.

WENDT, K. S., SCHALL, I., HUBER, R., BUCKEL, W. & JACOB, U. 2003. Crystal structure of the carboxyltransferase subunit of the bacterial sodium ion pump glutaconyl-coenzyme A decarboxylase. *The EMBO Journal,* 22**,** 3493-3502.

WHITE, M. D., PAYNE, K. A., FISHER, K., MARSHALL, S. A., PARKER, D., RATTRAY, N. J., TRIVEDI, D. K., GOODACRE, R., RIGBY, S. E., SCRUTTON, N. S., HAY, S. & LEYS, D. 2015. UbiX is a flavin prenyltransferase required for bacterial ubiquinone biosynthesis. *Nature,* 522**,** 502-506.

WIBERG, K. B. & SAEGEBARTH, K. A. 1957. The mechanisms of permanganate oxidation. IV. Hydroxylation of olefins and related reactions. *J Am Chem Soc,* 79**,** 2822-2824.

WILLIAMS, M. 2006. *The Merck Index: an Encyclopedia of Chemicals, Drugs, and Biologicals. 14th Edition.* , Wiley Subscription Services, Inc., A Wiley Company.

WILLKE, T. & VORLOP, K. D. 2001. Biotechnological production of itaconic acid. *Appl Microbiol Biotechnol,* 56**,** 289-295.

WONG, K. F., SELZER, T., BENKOVIC, S. J. & HAMMES-SCHIFFER, S. 2005. Impact of distal mutations on the network of coupled motions correlated to hydride transfer in dihydrofolate reductase. *Proc Natl Acad Sci U S A,* 102**,** 6807-6812.

XIAO, H., BAO, Z. & ZHAO, H. 2015. High throughput screening and selection methods for directed enzyme Evolution. *Ind Eng Chem Res,* 54**,** 4011-4020.

XU, H., ZHANG, Y., GUO, X., REN, S., STAEMPFLI, A. A., CHIAO, J., JIANG, W. & ZHAO, G. 2004. Isoleucine biosynthesis i*n Leptospira interrogans* serotype lai strain 56601 proceeds via a threonine-independent pathway. *J Bacteriol,* 186**,** 5400-5409.

YAHIRO, K., SHIBATA, S., JIA, S. R., PARK, Y. & OKABE, M. 1997. Efficient itaconic acid production from raw corn starch. *J Ferment Bioeng,* 84**,** 375-377.

YANG, G. & WITHERS, S. G. 2009. Ultrahigh-throughput FACS-based screening for directed enzyme evolution. *ChemBioChem,* 10**,** 2704-2715.

YANG, X. 2004. Processes for production of methyl methacrylate and their technology-economic comparison. *Chemical Industry and Engineering Progress,* 23**,** 506-510.

YAWS, C. L. 2003. Yaws' Handbook of Thermodynamic and Physical Properties of Chemical Compounds. Knovel.

YI, D., DEVAMANI, T., ABDOUL-ZABAR, J., CHARMANTRAY, F., HELAINE, V., HECQUET, L. & FESSNER, W.-D. 2012. A pH-based high-throughput assay for transketolase: Fingerprinting of substrate tolerance and quantitative Kinetics. *ChemBioChem,* 13**,** 2290-2300.

YUAN, J. & CHING, C. B. 2014. Combinatorial engineering of mevalonate pathway for improved amorpha-4,11-diene production in budding yeast. *Biotechnol Bioeng,* 111**,** 608-617.

ZHANG, J.-H., CHUNG, T. D. Y. & OLDENBURG, K. R. 1999. A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *J Biomol Screen,* 4**,** 67-73.

ZHANG, P., XU, C. & ZHU, H. 2002. New process of recycling methyl methacrylate from waste polymethyl methacrylate. *Chemical Industry and Engineering Progress,* 21**,** 753-755.

ZHANG, Y. 2008. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics,* 9**,** 40.

ZHENG, H. & REETZ, M. T. 2010. Manipulating the stereoselectivity of limonene epoxide hydrolase by directed evolution based on iterative saturation mutagenesis. *J Am Chem Soc,* 132**,** 15744-15751.

ZHENG, L., BAUMANN, U. & REYMOND, J. L. 2004. An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucleic Acids Res,* 32**,** e115.

ZHENG, Y., LIU, Q., LI, L., QIN, W., YANG, J., ZHANG, H., JIANG, X., CHENG, T., LIU, W., XU, X. & XIAN, M. 2013. Metabolic engineering of *Escherichia coli* for high-specificity production of isoprenol and prenol as next generation of biofuels. *Biotechnol Biofuels,* 6**,** 57.

ZHOU, X., CHEN, C., WANG, Y., XU, Z., DUAN, J., HAO, Y. & SMAILL, S. 2013. Soil extractable carbon and nitrogen, microbial biomass and microbial metabolic activity in response to warming and increased precipitation in a semiarid Inner Mongolian grassland. *Geoderma,* 206**,** 24-31.

ZHU, F., ZHONG, X., HU, M., LU, L., DENG, Z. & LIU, T. 2014. In vitro reconstitution of mevalonate pathway and targeted engineering of farnesene overproduction in *Escherichia coli*. *Biotechnol Bioeng,* 111**,** 1396-1405.

9 Appendices

9.1 System for overexpression of target enzymes

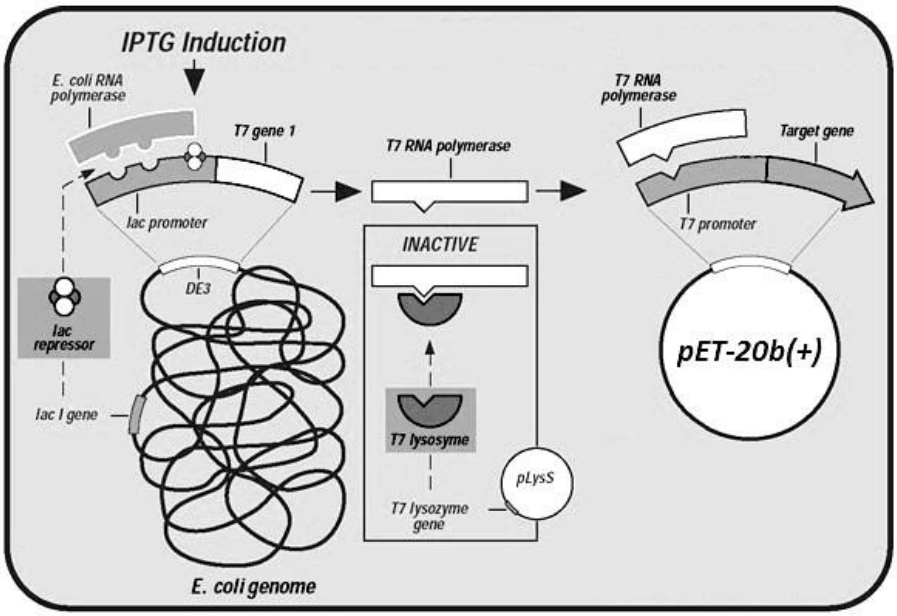


Figure 9.1: System for overexpression of target enzymes
The BL21(DE3)pLysS strain contains a chromosomal gene for expression of the T7 RNA polymerase under control of the *lac* promoter. The *lac* promoter is repressed in the presence of a *lac* repressor, also encoded by a chromosomal gene (*Lac I*). Upon addition of isopropyl β-D-1-thiogalactopyranoside (IPTG), the *lac* promoter is de-repressed and expression of T7 RNA polymerase begins. The gene encoding for the enzyme of interest, appositely inserted in pET-20b(+), is under control of a T7 promoter, a very strong promoter. Once the T7 RNA polymerase is present in the cells, it starts to transcribe the gene of interest and the expression of the desired enzyme begins. *E. coli* BL21(DE3)pLysS also contains an additional plasmid, pLysS, that produces the T7 lysozyme. The T7 lysozyme is a protein that inhibits the T7 RNA polymerase (Moffatt and Studier, 1987; Studier, 1991). This prevents that minimum concentrations of T7 RNA polymerase, given by leaky expression of the *lac* promoter even in the absence of IPTG, trigger the expression of the enzyme of interest. With this system, expression of the desired enzyme occurs only in the presence of IPTG. Finally, the T7 lysozyme also degrades the peptidoglycan layer of the *E. coli* cell wall (Inouye *et al.*, 1973) and, therefore, it facilitates lysis of cells for enzyme preparation.

9.2 Gene and amino acid sequences

S. cerevisiae PAD1

Gene sequence

(729 bp, NCBI Accession Number: NM_001180846)

ATGCTCCTATTTCCAAGAAGAACTAATATAGCCTTTTTCAAAACAACAGGCATTTTTGCTAATTTTCC
TTTGCTAGGTAGAACCATTACAACTTCACCATCTTTCCTTACACATAAACTGTCAAAGGAAGTAACCAGG
GCATCAACTTCGCCTCCAAGACCAAAGAGAATTGTTGTCGCAATTACTGGTGCGACTGGTGTTGCACTGG
GAATCAGACTTCTACAAGTGCTAAAAGAGTTGAGCGTAGAAACCCATTTGGTGATTTCAAAATGGGGTGC
AGCAACAATGAAATATGAAACAGATTGGGAACCGCATGACGTGGCGGCCTTGGCAACCAAGACATACTCT
GTTCGTGATGTTTCTGCATGCATTTCGTCCGGATCTTTCCAGCATGATGGTATGATTGTTGTGCCCTGTT
CCATGAAATCACTAGCTGCTATTAGAATCGGTTTTACAGAGGATTTAATTACAAGAGCTGCCGATGTTTC
GATTAAAGAGAATCGTAAGTTACTACTGGTTACTCGGGAAACCCCTTTATCTTCCATCCATCTTGAAAAC
ATGTTGTCTTTATGCAGGGCAGGTGTTATAATTTTTCCTCCGGTACCTGCGTTTTATACAAGACCCAAGA
GCCTTCATGACCTATTAGAACAAAGTGTTGGCAGGATCCTAGACTGCTTTGGCATCCACGCTGACACTTT
TCCTCGTTGGGAAGGAATAAAAAGCAAGTAA

Amino acid sequence

(242 aa, NCBI accession number NP_010827)

MLLFPRRTNIAFFKTTGIFANFPLLGRTITTSPSFLTHKLSKEVTRASTSPPRPKRIVVAITGATGVA
LGIRLLQVLKELSVETHLVISKWGAATMKYETDWEPHDVAALATKTYSVRDVSACISSGSFQHDGMIVVP
CSMKSLAAIRIGFTEDLITRAADVSIKENRKLLLVTRETPLSSIHLENMLSLCRAGVIIFPPVPAFYTRP
KSLHDLLEQSVGRILDCFGIHADTFPRWEGIKSK

S. cerevisiae FDC1

Gene sequence

(1512 bp, NCBI Accession Number: NM_001180847)

ATGAGGAAGCTAAATCCAGCTTTAGAATTTAGAGACTTTATCCAGGTCTTAAAAGATGAAGATGACTT
AATCGAAATTACCGAAGAGATTGATCCAAATCTCGAAGTAGGTGCAATTATGAGGAAGGCCTATGAATCC
CACTTACCAGCCCCGTTATTTAAAAATCTCAAAGGTGCTTCGAAGGATCTTTTCAGCATTTTAGGTTGCC
CAGCCGGTTTGAGAAGTAAGGAGAAAGGAGATCATGGTAGAATTGCCCATCATCTGGGGCTCGACCCAAA
AACAACTATCAAGGAAATCATAGATTATTTGCTGGAGTGTAAGGAGAAGGAACCTCTCCCCCCAATCACT
GTTCCTGTGTCATCTGCACCTTGTAAAACACATATACTTTCTGAAGAAAAAATACATCTACAAAGCCTGC
CAACACCATATCTACATGTTTCAGACGGTGGCAAGTACTTACAAACGTACGGAATGTGGATTCTTCAAAC
TCCAGATAAAAAATGGACTAATTGGTCAATTGCTAGAGGTATGGTTGTAGATGACAAGCATATCACTGGT
CTGGTAATTAAACCACAACATATTAGACAAATTGCTGACTCTTGGGCAGCAATTGGAAAAGCAAATGAAA
TTCCTTTCGCGTTATGTTTTGGCGTTCCCCCAGCAGCTATTTTAGTTAGTTCCATGCCAATTCCTGAAGG
TGTTTCTGAATCGGATTATGTTGGCGCAATCTTGGGTGAGTCGGTTCCAGTAGTAAAATGTGAGACCAAC
GATTTAATGGTTCCTGCAACGAGTGAGATGGTATTTGAGGGTACTTTGTCCTTAACAGATACACATCTGG
AAGGCCCATTTGGTGAGATGCATGGATATGTTTTCAAAAGCCAAGGTCATCCTTGTCCATTGTACACTGT
CAAGGCTATGAGTTACAGAGACAATGCTATTCTACCTGTTTCGAACCCCGGTCTTTGTACGGATGAGACA
CATACCTTGATTGGTTCACTAGTGGCTACTGAGGCCAAGGAGCTGGCTATTGAATCTGGCTTGCCAATTC
TGGATGCCTTTATGCCTTATGAGGCTCAGGCTCTTTGGCTTATCTTAAAGGTGGATTTGAAAGGGCTGCA
AGCATTGAAGACAACGCCTGAAGAATTTTGTAAGAAGGTAGGTGATATTTACTTTAGGACAAAAGTTGGT
TTTATAGTCCATGAAATAATTTTGGTGGCAGATGATATCGACATATTTAACTTCAAAGAAGTCATCTGGG
CCTACGTTACAAGACATACACCTGTTGCAGATCAGATGGCTTTTGATGATGTCACTTCTTTTCCTTTGGC
TCCCTTTGTTTCGCAGTCATCCAGAAGTAAGACTATGAAAGGTGGAAAGTGCGTTACTAATTGCATATTT
AGACAGCAATATGAGCGCAGTTTTGACTACATAACTTGTAATTTTGAAAAGGGATATCCAAAAGGATTAG
TTGACAAAGTAAATGAAAATTGGAAAAGGTACGGATATAAATAA

Amino acid sequence

(503 aa, NCBI accession number NP_010828)

MRKLNPALEFRDFIQVLKDEDDLIEITEEIDPNLEVGAIMRKAYESHLPAPLFKNLKGASKDLFSIL
GCPAGLRSKEKGDHGRIAHHLGLDPKTTIKEIIDYLLECKEKEPLPPITVPVSSAPCKTHILSEEKIHL
QSLPTPYLHVSDGGKYLQTYGMWILQTPDKKWTNWSIARGMVVDDKHITGLVIKPQHIRQIADSWAAIG
KANEIPFALCFGVPPAAILVSSMPIPEGVSESDYVGAILGESVPVVKCETNDLMVPATSEMVFEGTLSL
TDTHLEGPFGEMHGYVFKSQGHPCPLYTVKAMSYRDNAILPVSNPGLCTDETHTLIGSLVATEAKELAI
ESGLPILDAFMPYEAQALWLILKVDLKGLQALKTTPEEFCKKVGDIYFRTKVGFIVHEIILVADDIDIF
NFKEVIWAYVTRHTPVADQMAFDDVTSFPLAPFVSQSSRSKTMKGGKCVTNCIFRQQYERSFDYITCNF
EKGYPKGLVDKVNENWKRYGYK

A. niger PadA1

Gene sequence

(684 bp, NCBI Accession Number: EF215454)

ATGTTCAACTCACTTCTATCCGGTACTACTACACCAAACTCCGGCCGTGCATCTCCTCCCGCCAGCGA
AATGCCCATCGATAATGATCACGTGGCCGTTGCCCGTCCAGCTCCCCGCCGCCGCCGCATTGTAGTAGCC
ATGACGGGTGCCACTGGAGCCATGCTCGGCATCAAAGTCCTAATTGCTCTGCGCCGTCTAAATGTGGAGA
CACACCTGGTGATGAGTAAATGGGCGGAGGCTACGATCAAATACGAGACTGACTACCATCCCTCAAACGT
GCGAGCGCTGGCCGACTACGTGCACAACATCAATGACATGGCCGCCCCAGTATCCAGCGGCTCATTCCGC
GCGGACGGAATGATTGTGGTACCGTGCAGCATGAAAACATTGGCTGCTATCCACTCGGGCTTTTGCGACG
ATCTCATTTCAAGGACAGCAGATGTGATGCTCAAGGAGCGCAGGCGGTTGGTGCTAGTAGCGCGGGAGAC
GCCATTGAGCGAGATCCATCTGCGAAACATGTTGGAGGTTACACGCGCTGGGGCAGTCATCTTCCCCCCA
GTACCGGCGTTCTACATCAAGGCCGGAAGTATCGAGGACCTCATCGACCAGAGTGTTGGACGAATGTTGG
ATTTATTTGACCTCGACACGGGGGATTTTGAACGTTGGAATGGATGGGAAAAATAA

Amino acid sequence

(227 aa, NCBI accession number ABN13117)

MFNSLLSGTTTPNSGRASPPASEMPIDNDHVAVARPAPRRRRIVVAMTGATGAMLGIKVLIALRRLN
VETHLVMSKWAEATIKYETDYHPSNVRALADYVHNINDMAAPVSSGSFRADGMIVVPCSMKTLAAIHSG
FCDDLISRTADVMLKERRRLVLVARETPLSEIHLRNMLEVTRAGAVIFPPVPAFYIKAGSIEDLIDQSV
GRMLDLFDLDTGDFERWNGWEK

A. niger OhbA1

Gene sequence

(1503 bp, NCBI Accession Number: XM_001390497)

ATGTCTGCGCAACCTGCTCACCTGTGTTTCCGCTCCTTCGTCGAAGCCCTCAAGGTCGACAACGACCT
TGTTGAAATCAATACCCCAATTGACCCCAATCTCGAAGCTGCTGCTATTACCCGCCGAGTATGTGAGACC
AACGACAAGGCTCCTTTATTCAACAACCTCATCGGCATGAAAAATGGCCTCTTCCGTATACTTGGGGCTC
CTGGCTCTCTCAGGAAGTCGTCTGCTGATCGCTACGGCCGCCTTGCTCGTCACCTAGCCCTCCCACCTAC
GGCCTCAATGCGTGAGATTCTCGATAAGATGCTCTCCGCCAGCGATATGCCTCCCATCCCTCCGACCATT
GTTCCCACCGGGCCATGCAAGGAGAACAGCTTAGATGACTCTGAATTCGACCTTACCGAACTCCCCGTTC
CTCTTATTCACAAATCGGATGGTGGTAAATACATCCAAACCTATGGCATGCACATTGTGCAGTCTCCGGA
TGGAACCTGGACCAACTGGTCTATTGCCCGTGCGATGGTCCATGACAAGAACCATCTGACCGGCCTGGTT
ATTCCCCCTCAGCACATCTGGCAGATTCACCAGATGTGGAAGAAGGAAGGCCGCAGTGACGTTCCCTGGG
CTTTGGCCTTTGGTGTCCCACCCGCTGCCATTATGGCCTCTAGCATGCCTATTCCCGATGGTGTCACCGA
AGCTGGGTACGTGGGAGCTATGACGGGATCCTCCCTGGAGCTTGTTAAATGTGATACGAACGATCTATAT
GTCCCCGCTACCTCAGAAATCGTTCTCGAGGGCACACTCTCTATCAGCGAGACAGGCCCAGAGGGACCTT
TCGGTGAGATGCATGGTTACATCTTCCCCGGGGATACTCACCTCGGCGCCAAATACAAGGTTAACCGGAT
CACCTACCGCAACAACGCCATCATGCCCATGTCTTCTTGTGGCCGCTTGACGGATGAAAC**(INTRON)**G
CACACCATGATCGGCTCTCTGGCTGCGGCGGAGATCCGTAAGCTCTGCCAGCAGAATGACCTCCCTATCA
CTGATGCCTTCGCTCCTTTCGAGTCTCAAGTTACCTGGGTTGCTCTGCGGGTCGATACTGAGAAGCTACG
TGCCATGAAGACAACGTCTGAGGGATTCCGCAAGAGAGTGGGAGACGTCGTCTTCAACCACAAGGCCGGA
TACACCATTCATCGTCTGGTGTTGGTCGGTGACGACATTGATGTCTATGAAGGAAAGGATGTGCTCTGGG
CGTTCTCCACCCGTTGCCGTCCTGGTATGGACGAGACTTTGTTTGAGGATGTTCGTGGGTTCCCCTTGAT
TCCGTATATGGGACACGGGAATGGGCCCGCCCACCGCGGCGGAAAGGTTGTGTCCGACGCTCTTATGCCG
ACTGAGTACACCACTGGTCGCAACTGGGAGGCTGCTGACTTCAACCAATCTTATCCCGAGGATCTGAAGC
AGAAGGTGTTGGACAACTGGACGAAGATGGGTTTCAGCAACTAA

Amino acid sequence

(500 aa, NCBI accession number: XP_001390534)

MSAQPAHLCFRSFVEALKVDNDLVEINTPIDPNLEAAAITRRVCETNDKAPLFNNLIGMKNGLFRILG
APGSLRKSSADRYGRLARHLALPPTASMREILDKMLSASDMPPIPPTIVPTGPCKENSLDDSEFDLTELP
VPLIHKSDGGKYIQTYGMHIVQSPDGTWTNWSIARAMVHDKNHLTGLVIPPQHIWQIHQMWKKEGRSDVP
WALAFGVPPAAIMASSMPIPDGVTEAGYVGAMTGSSLELVKCDTNDLYVPATSEIVLEGTLSISETGPEG
PFGEMHGYIFPGDTHLGAKYKVNRITYRNNAIMPMSSCGRLTDETHTMIGSLAAAEIRKLCQQNDLPITD
AFAPFESQVTWVALRVDTEKLRAMKTTSEGFRKRVGDVVFNHKAGYTIHRLVLVGDDIDVYEGKDVLWAF
STRCRPGMDETLFEDVRGFPLIPYMGHGNGPAHRGGKVVSDALMPTEYTTGRNWEAADFNQSYPEDLKQK
VLDNWTKMGFSN

A. terreus CadA1

Gene sequence

(1473 bp without intron, NCBI Accession Number: AB326105)

ATGACCAAACAATCTGCGGACAGCAACGCAAAGTCAGGAGTTACGTCCGAAATATGTCATTGGGCATC
CAACCTGGCCACTGACGACATCCCTTCGGACGTATTAGAAAGAGCAAAATACCTTATTCTCGACGGTATT
GCATGTGCCTGGGTTGGTGCAAGAGTGCCTTGGTCAGAGAAGTATGTTCAGGCAACGATGAGCTTTGAGC
CGCCGGGGGCCTGCAGGGTGATTGGATATGGACAG**(INTRON)**AAACTGGGGCCTGTTGCAGCAGCCAT
GACCAATTCCGCTTTCATACAGGCTACGGAGCTTGAC

ACTACCACAGCGAAGCCCCCCTACACTCTGCAAGCATTGTCCTTCCTGCGGTCTTTGCAGCAAGTGAG
GTCTTAGCCGAGCAGGGCAAAACAATTTCCGGTATAGATGTTATTCTAGCCGCCATTGTGGGGTTTGAAT
CTGGCCCACGGATCGGCAAAGCAATCTACGGATCGGACCTCTTGAACAACGGCTGGCATTGTGGAGCTGT
GTATGGCGCTCCAGCCGGTGCGCTGGCCACAGGAAAGCTCCTCGGTCTAACTCCAGACTCCATGGAAGAT
GCTCTCGGAATTGCGTGCACGCAAGCCTGTGGTTTAATGTCGGCGCAATACGGAGGCATGGTAAAGCGTG
TGCAACACGGATTCGCAGCGCGTAATGGTCTTCTTGGGGGACTGTTGGCCCATGGTGGGTACGAGGCAAT
GAAAGGTGTCCTGGAGAGATCTTACGGCGGTTTCCTCAAGATGTTCACCAAGGGCAACGGCAGAGAGCCT
CCCTACAAAGAGGAGGAAGTGGTGGCTGGTCTCGGTTCATTCTGGCATACCTTTACTATTCGCATCAAGC
TCTATGCCTGCTGCGGACTTGTCCATGGTCCAGTCGAGGCTATCGAAAACCTTCAGGGGAGATACCCCGA
GCTCTTGAATAGAGCCAACCTCAGCAACATTCGCCATGTTCATGTACAGCTTTCAACGGCCTCGAACAGT
CACTGTGGATGGATACCAGAGGAGAGACCCATCAGTTCAATCGCAGGGCAGATGAGTGTCGCATACATTC
TCGCCGTCCAGCTGGTCGACCAGCAATGTCTTTTGTCCCAGTTTTCTGAGTTTGATGACAACCTGGAGAG
GCCAGAAGTTTGGGATCTGGCCAGGAAGGTTACTTCATCTCAAAGCGAAGAGTTTGATCAAGACGGCAAC
TGTCTCAGTGCGGGTCGCGTGAGGATTGAGTTCAACGATGGTTCTTCTATTACGGAAAGTGTCGAGAAGC
CTCTTGGTGTCAAAGAGCCCATGCCAAACGAACGGATTCTCCACAAATACCGAACCCTTGCTGGTAGCGT
GACGGACGAATCCCGGGTGAAAGAGATTGAGGATCTTGTCCTCGGCCTGGACAGGCTCACCGACATTAGC
CCATTGCTGGAGCTGCTGAATTGCCCCGTGAAATCGCCACTGGTATAA

Amino acid sequence

(490 aa, NCBI accession number: BAG49047)

MTKQSADSNAKSGVTSEICHWASNLATDDIPSDVLERAKYLILDGIACAWVGARVPWSEKYVQATMSF
EPPGACRVIGYGQKLGPVAAAMTNSAFIQATELDDYHSEAPLHSASIVLPAVFAASEVLAEQGKTISGID
VILAAIVGFESGPRIGKAIYGSDLLNNGWHCGAVYGAPAGALATGKLLGLTPDSMEDALGIACTQACGLM
SAQYGGMVKRVQHGFAARNGLLGGLLAHGGYEAMKGVLERSYGGFLKMFTKGNGREPPYKEEEVVAGLGS
FWHTFTIRIKLYACCGLVHGPVEAIENLQGRYPELLNRANLSNIRHVHVQLSTASNSHCGWIPEERPISS
IAGQMSVAYILAVQLVDQQCLLSQFSEFDDNLERPEVWDLARKVTSSQSEEFDQDGNCLSAGRVRIEFND
GSSITESVEKPLGVKEPMPNERILHKYRTLAGSVTDESRVKEIEDLVLGLDRLTDISPLLELLNCPVKSP
LV

P. fluorescens ACMSD

Gene sequence

(1005 bp, optimized for expression in *E. coli*, original sequence NCBI Accession Number: AB088043)

ATGAAGAAGCCGCGCATTGACATGCATTCCCACTTCTTCCCGCGTATCAGCGAACAGGAGGCGGCGAA
ATTCGATGCTAACCACGCTCCGTGGCTGCAGGTGTCCGCTAAAGGTGACACCGGTTCTATTATGATGGGT
AAAAACAACTTCCGTCCGGTTTACCAGGCACTGTGGGACCCGGCGTTCCGTATCGAAGAAATGGATGCGC
AGGGTGTTGATGTTCAGGTTACCTGTGCAACCCCGGTTATGTTCGGTTACACTTGGGAAGCAAACAAAGC
AGCGCAGTGGGCGGAACGTATGAACGACTTCGCGCTGGAGTTCGCGGCTCACAACCCGCAGCGTATTAAA
GTTCTGGCGCAGGTTCCGCTGCAGGATCTGGACCTGGCTTGCAAAGAAGCGTCTCGTGCGGTTGCGGCAG
GTCACCTGGGTATTCAGATCGGCAACCATCTGGGCGACAAAGACCTGGACGACGCGACCCTGGAAGCGTT
TCTGACCCATTGCGCGAACGAGGACATCCCGATCCTGGTGCATCCGTGGGACATGATGGGCGGTCAGCGC
ATGAAAAAGTGGATGCTGCCGTGGCTGGTTGCGATGCCGGCTGAAACTCAGCTGGCGATTCTGTCCCTGA
TTCTGAGCGGTGCGTTCGAGCGTATCCCGAAATCTCTGAAAATTTGTTTCGGTCATGGTGGTGGTTCTTT
TGCGTTCCTGCTGGGCCGTGTGGATAACGCATGGCGTCACCGTGATATCGTTCGTGAAGATTGCCCGCGT
CCGCCGTCTGAGTACGTTGACCGTTTTTTCGTGGACAGCGCGGTTTTCAACCCGGGTGCACTGGAACTGC
TGGTTTCTGTTATGGGTGAAGACCGTGTGATGCTGGGTTCCGATTACCCGTTCCCGCTGGGTGAGCAGAA
GATCGGTGGCCTGGTTCTGTCTTCTAACCTGGGCGAGTCTGCTAAGGATAAGATCATTTCTGGTAACGCG
TCTAAATTCTTCAACATTAACGTTTAA

Amino acid sequence

(334 aa, NCBI accession number: BAC65312)

MKKPRIDMHSHFFPRISEQEAAKFDANHAPWLQVSAKGDTGSIMMGKNNFRPVYQALWDPAFRIEEMD
AQGVDVQVTCATPVMFGYTWEANKAAQWAERMNDFALEFAAHNPQRIKVLAQVPLQDLDLACKEASRAVA
AGHLGIQIGNHLGDKDLDDATLEAFLTHCANEDIPILVHPWDMMGGQRMKKWMLPWLVAMPAETQLAILS
LILSGAFERIPKSLKICFGHGGGSFAFLLGRVDNAWRHRDIVREDCPRPPSEYVDRFFVDSAVFNPGALE
LLVSVMGEDRVMLGSDYPFPLGEQKIGGLVLSSNLGESAKDKIISGNASKFFNINV

P. fluorescens HAO

Gene sequence

(558 bp, optimized for expression in *E. coli*, original sequence NCBI Accession Number: AB088043)

ATGATGTTTACATTCGGCAAACCATTGAATTTTCAGCGTTGGCTCGATGACCATTCTGACTTGTTGAG
ACCACCGGTTGGAAATCAACAAGTTTGGCAAGACTCCGATTTTATTGTGACAGTGGTCGGTGGGCCAAAT
TTTCGCACAGATTTTCATGATGATCCGATGGAAGAATTTTTCTATCAGTTCAAGGGTAATGCTTACTTGA
ACATCATGGACCGAGGTCAGATGGATCGTGTCGAGTTGAAGGAAGGCGATATATTTCTTCTGCCACCGCA
TTTGCGCCACTCGCCACAACGACCTGAAGCAGGAAGTCGCTGCCTGGTGATCGAGCGTCAACGGCCTAAA
GGTATGCTTGATGGTTTTGAGTGGTATTGCCTAAGCTGTAATGGCCTGGTTTATCGGGTGGATGTTCAGC
TCAATAGTATCGTCACCGATTTGCCGCCATTGTTTGATATTTTTTATGGCAACGTCGGTTTGCGCAAGTG
CCCTCAATGTGGGCAAGTCCATCCAGGTAAAGCAGCGATCGAAGCGGTGGCGCGGGGAGATCAGCCATGA

Amino acid sequence

(185 aa, NCBI accession number**:** BAC65311)

MMFTFGKPLNFQRWLDDHSDLLRPPVGNQQVWQDSDFIVTVVGGPNFRTDFHDDPMEEFFYQFKGNAY
LNIMDRGQMDRVELKEGDIFLLPPHLRHSPQRPEAGSRCLVIERQRPKGMLDGFEWYCLSCNGLVYRVDV
QLNSIVTDLPPLFDIFYGNVGLRKCPQCGQVHPGKAAIEAVARGDQP

S. cerevisiae MVD

Gene sequence

(1191 bp, optimized for expression in *E. coli*, original sequence NCBI Accession Number: NM_001183220)

ATGACCGTTTACACCGCGTCTGTGACTGCTCCGGTGAACATCGCAACTCTGAAATACTGGGGTAAGCG
TGATACTAAACTGAACCTGCCGACCAACAGCTCTATCAGCGTTACTCTGAGCCAGGACGATCTGCGTACC
CTGACCTCTGCAGCTACTGCGCCGGAATTCGAGCGTGACACCCTGTGGCTGAACGGTGAGCCGCACTCTA
TCGACAACGAACGTACCCAGAACTGCCTGCGTGACCTGCGCCAGCTGCGTAAAGAGATGGAATCCAAGGA
TGCGAGCCTGCCGACTCTGTCTCAGTGGAAACTGCACATCGTTTCTGAGAACAACTTCCCGACCGCAGCG
GGCCTGGCGTCTTCTGCTGCGGGTTTCGCGGCACTGGTTAGCGCGATTGCGAAACTGTACCAGCTGCCGC
AGTCTACCTCTGAGATTTCTCGTATCGCGCGCAAGGGCTCTGGCTCTGCGTGCCGTTCTCTGTTTGGTGG
CTATGTTGCGTGGGAGATGGGCAAAGCTGAGGACGGTCACGACTCTATGGCGGTTCAGATCGCGGACTCT
TCTGACTGGCCGCAGATGAAAGCGTGTGTTCTGGTTGTGAGCGACATCAAGAAGGACGTTTCTTCTACCC
AGGGTATGCAGCTGACCGTGGCGACCAGCGAACTGTTTAAGGAGCGTATCGAACACGTTGTTCCGAAGCG
TTTTGAGGTTATGCGTAAGGCTATTGTTGAAAAGGACTTCGCAACCTTCGCTAAAGAAACCATGATGGAC
TCCAACTCCTTTCACGCGACCTGCCTGGATTCTTTCCCGCCGATTTTCTACATGAACGACACCTCTAAAC
GTATCATTTCCTGGTGCCACACCATCAACCAGTTCTATGGTGAAACTATCGTTGCGTACACCTTCGACGC
GGGTCCGAACGCTGTTCTGTACTACCTGGCGGAAAACGAATCTAAGCTGTTTGCATTCATTTACAAACTG
TTCGGTTCTGTTCCGGGCTGGGACAAAAAGTTCACCACTGAACAGCTGGAGGCGTTTAACCACCAGTTTG
AGTCTTCTAACTTCACCGCGCGTGAGCTGGACCTGGAACTGCAGAAAGACGTTGCGCGTGTTATCCTGAC
CCAGGTTGGTTCTGGCCCGCAGGAAACCAACGAGTCCCTGATCGATGCGAAAACCGGTCTGCCGAAGGAA
TAA

Amino acid sequence

(396 aa, NCBI accession number*:* NP_014441)

MTVYTASVTAPVNIATLKYWGKRDTKLNLPTNSSISVTLSQDDLRTLTSAATAPEFERDTLWLNGEPH
SIDNERTQNCLRDLRQLRKEMESKDASLPTLSQWKLHIVSENNFPTAAGLASSAAGFAALVSAIAKLYQL
PQSTSEISRIARKGSGSACRSLFGGYVAWEMGKAEDGHDSMAVQIADSSDWPQMKACVLVVSDIKKDVSS
TQGMQLTVATSELFKERIEHVVPKRFEVMRKAIVEKDFATFAKETMMDSNSFHATCLDSFPPIFYMNDTS
KRIISWCHTINQFYGETIVAYTFDAGPNAVLYYLAENESKLFAFIYKLFGSVPGWDKKFTTEQLEAFNHQ
FESSNFTARELDLELQKDVARVILTQVGSGPQETNESLIDAKTGLPKE

P. torridus MVD

Gene sequence

(978 bp, optimized for expression in *E. coli*, original sequence NCBI Accession Number: NC_005877)

ATGGAGAACTATAATGTTAAAACCCGTGCATTTCCGACCATTGGTATTATTCTGCTGGGTGGCATTAG
CGACAAAAAAAACCGTATTCCGCTGCATACCACCGCAGGTATTGCATATACCGGCATCAATAACGATGTG
TACACCGAAACCAAACTGTATGTGAGCAAAGACGAAAAATGCTATATCGATGGCAAAGAAATCGATCTGA
ATAGCGATCGTAGCCCGAGCAAAGTGATCGATAAATTCAAACATGAAATCCTGATGCGTGTGAATCTGGA
TGATGAAAACAACCTGAGCATTGATAGCCGCAATTTTAACATTCTGAGCGGTAGCAGCGATAGCGGTGCA
GCAGCACTGGGTGAATGCATTGAAAGCATCTTCGAGTACAACATCAACATCTTCACCTTTGAAAATGATC
TGCAGCGTATTAGCGAAAGCGTTGGTCGTAGCCTGTATGGTGGTCTGACCGTTAATTATGCAAATGGTCG
TGAAAGCCTGACCGAACCGCTGCTGGAACCGGAAGCATTTAACAACTTTACCATCATCGGTGCCCATTTT
AACATTGATCGCAAACCGAGCAACGAAATCCACGAAAACATCATCAAACATGAGAACTATCGCGAACGTA
TTAAAAGCGCAGAGCGCAAAGCAAAAAAACTGGAAGAACTGAGCCGTAATGCCAACATTAAAGGCATTTT
TGAACTGGCAGAAAGCGATACCGTGGAATATCATAAAATGCTGCATGATGTGGGCGTTGATATTATCAAT
GACCGCATGGAAAATCTGATTGAACGCGTGAAAGAGATGAAAAACAACTTCTGGAACAGCTATATTGTTA
CCGGTGGTCCGAATGTTTTTGTGATCACCGAGAAAAAAGATGTGGATAAAGCCATGGAAGGTCTGAATGA
TCTGTGTGATGATATTCGTCTGCTGAAAGTTGCAGGTAAACCGCAGGTTATCAGCAAAAACTTCTAATGA

Amino acid sequence

(324 aa, NCBI accession number: YP_024134)

MENYNVKTRAFPTIGIILLGGISDKKNRIPLHTTAGIAYTGINNDVYTETKLYVSKDEKCYIDGKEID
LNSDRSPSKVIDKFKHEILMRVNLDDENNLSIDSRNFNILSGSSDSGAAALGECIESIFEYNINIFTFEN
DLQRISESVGRSLYGGLTVNYANGRESLTEPLLEPEAFNNFTIIGAHFNIDRKPSNEIHENIIKHENYRE
RIKSAERKAKKLEELSRNANIKGIFELAESDTVEYHKMLHDVGVDIINDRMENLIERVKEMKNNFWNSYI
VTGGPNVFVITEKKDVDKAMEGLNDLCDDIRLLKVAGKPQVISKNF

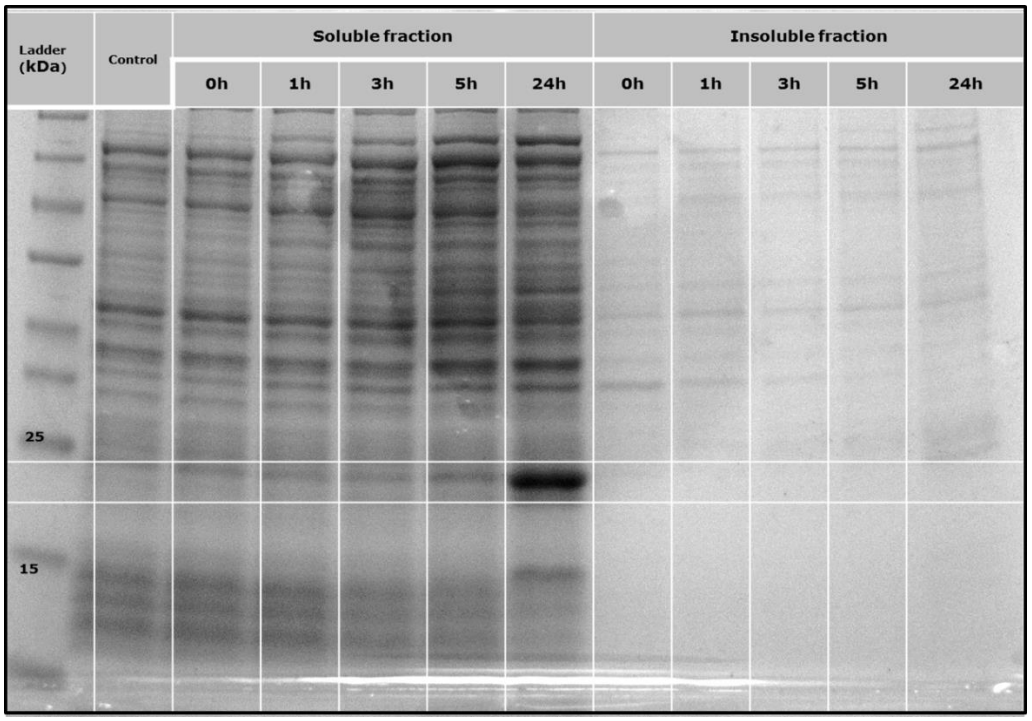
9.3 Optimisation of enzyme expression

Expression levels of the enzymes selected for a first screening for MAA formation were analysed by SDS-PAGE. The enzymes were: *S. cerevisiae* PAD1 and FDC1, *A. niger* PadA1 and OhbA1, *A. terreus* CadA1, *P. fluorescens* HAO and ACMSD, *S. cerevisiae* and *P. torridus* MVDs. After expression at different temperatures, both the soluble and the insoluble protein fractions were extracted and analysed by SDS-PAGE. The following figures show the gel results. Each lane of the gel was loaded with different samples as described below as labelled in the figures

- **Ladder:** mixture of proteins having defined molecular weights (indicated in kDa) which allowed mass evaluation of unknown proteins by comparing migration.
- **Control:** soluble protein fraction extracted from *E. coli* BL21(DE3)pLysS transformed with pET20b(+) not containing the genes of interest. This is a negative control.
- **Soluble:** soluble protein fraction extracted from *E. coli* BL21(DE3)pLysS expressing the enzymes of interest. Samples were taken at intervals (in hours) after IPTG induction, as indicated in the pictures.
- **Insoluble:** insoluble protein fraction extracted from *E. coli* BL21(DE3)pLysS expressing the enzymes of interest. Samples were taken at intervals (in hours) after IPTG induction, as indicated in the pictures.

The level of migration corresponding to the enzymes molecular weight (MW) is marked between white lines in the figures.

A



B

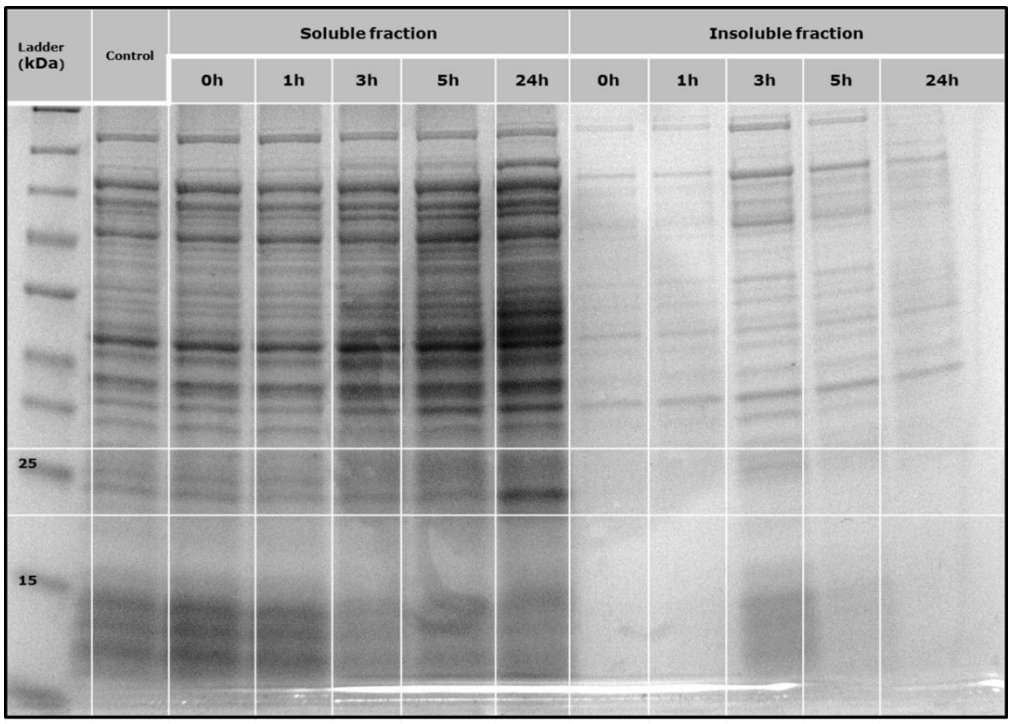
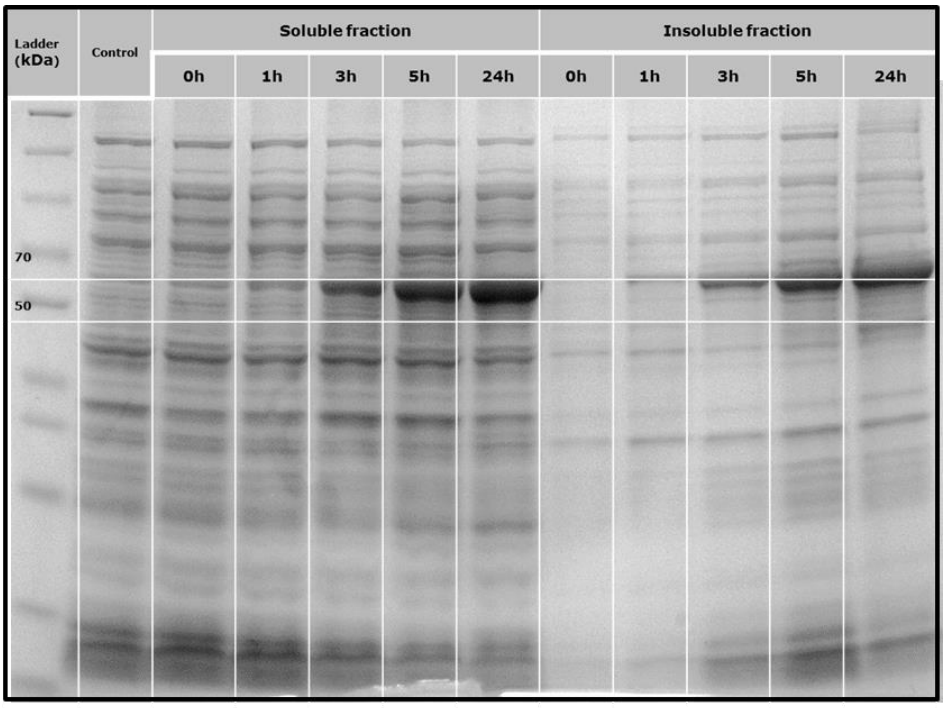


Figure 9.2: Expression of S. *cerevisiae* PAD1
Expression of the enzyme at 30˚C (A) and 37˚C (B).

A



B

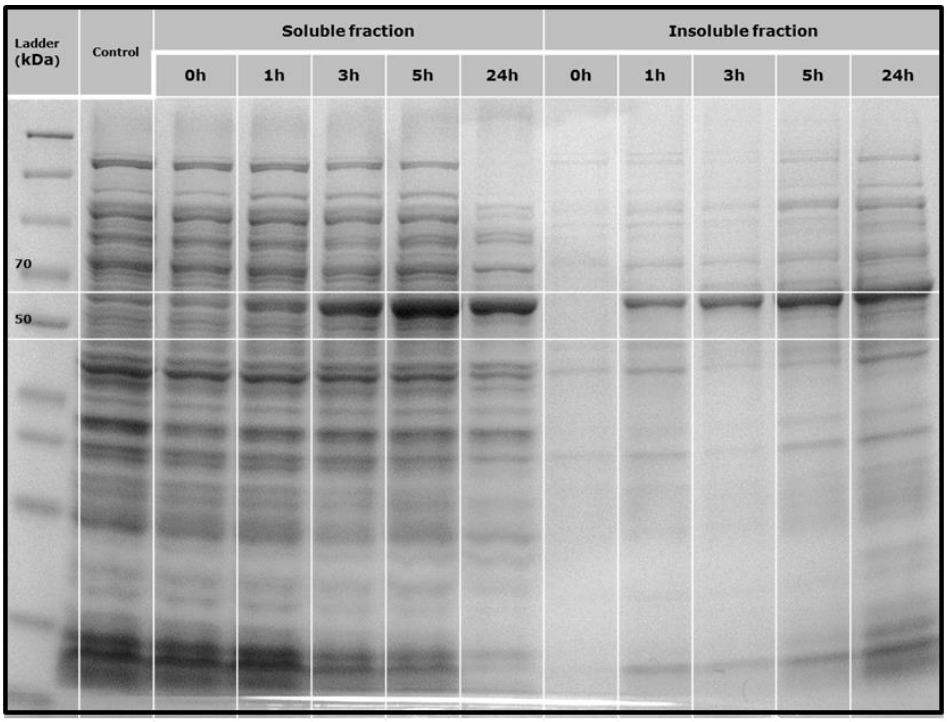
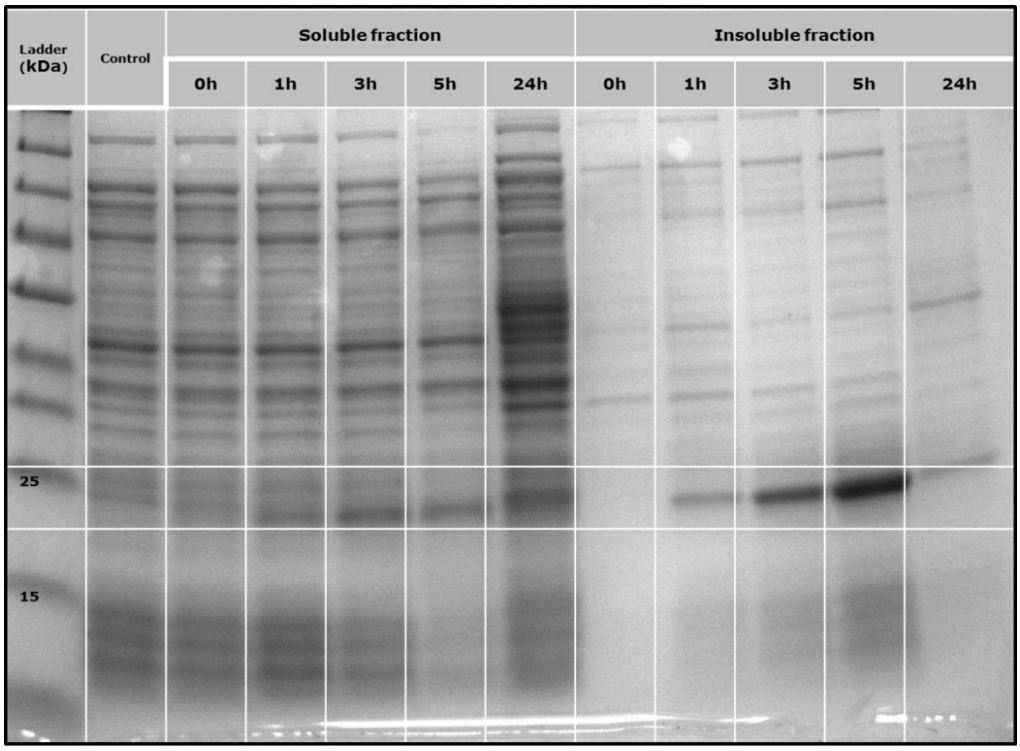


Figure 9.3: Expression of S. *cerevisiae* FDC1
Expression of the enzyme at 30˚C (A) and 37˚C (B).

A



B

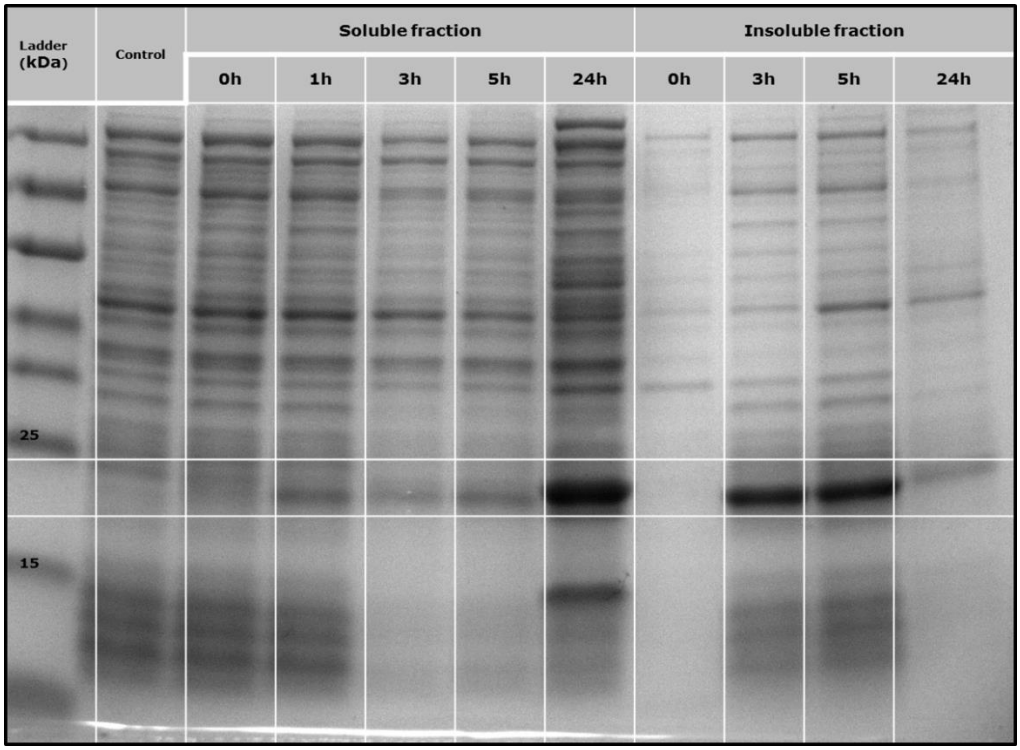
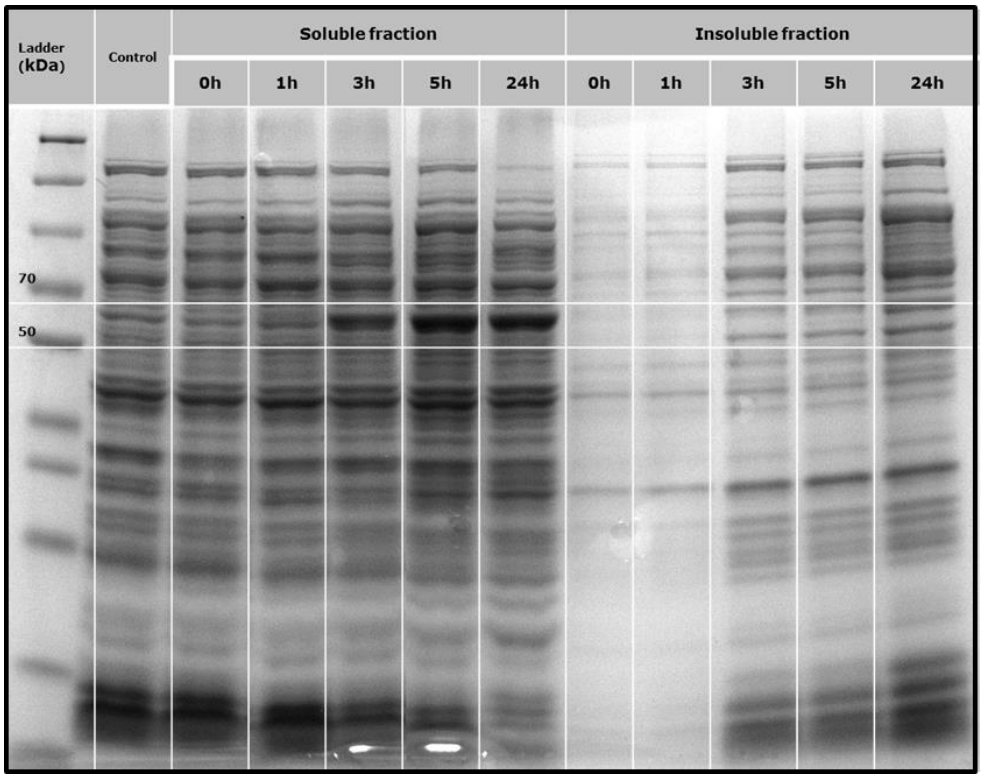


Figure 9.4: Expression of *A. niger* PadA1
Expression of the enzyme at 30˚C (A) and 37˚C (B).

A



B

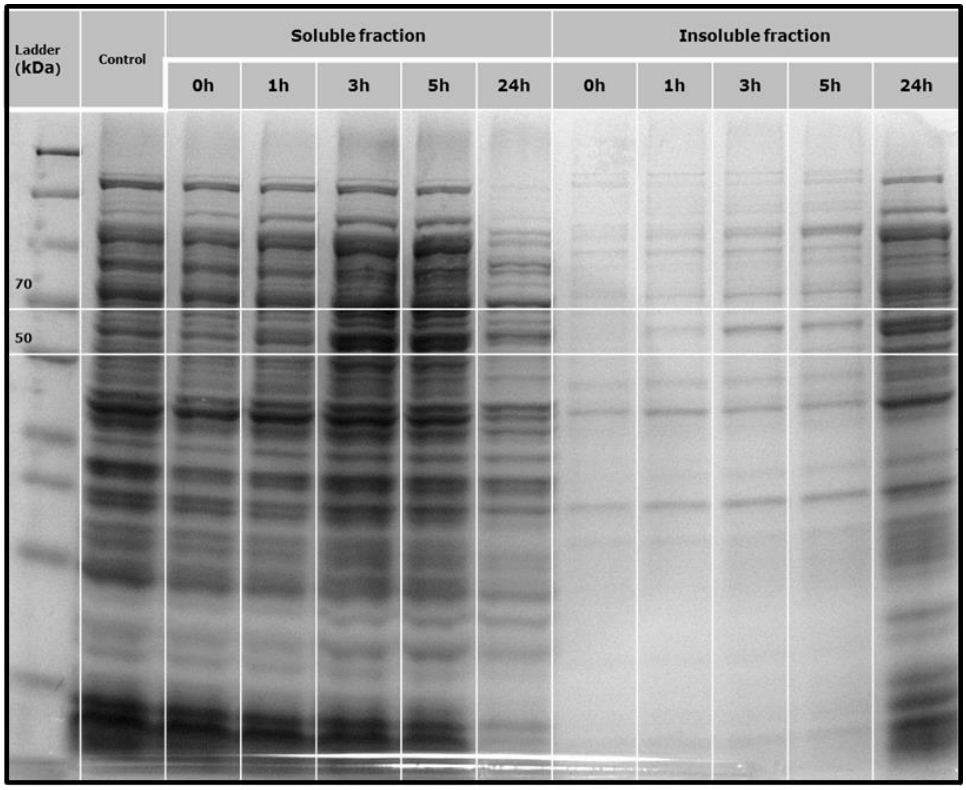
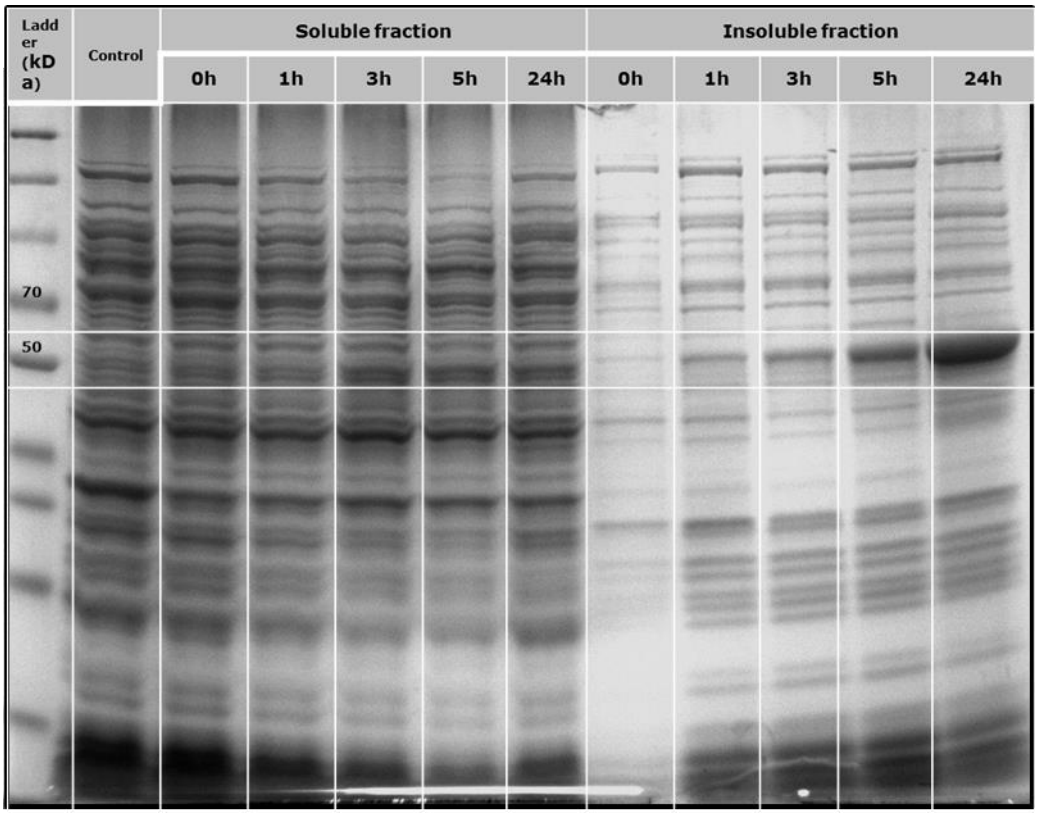
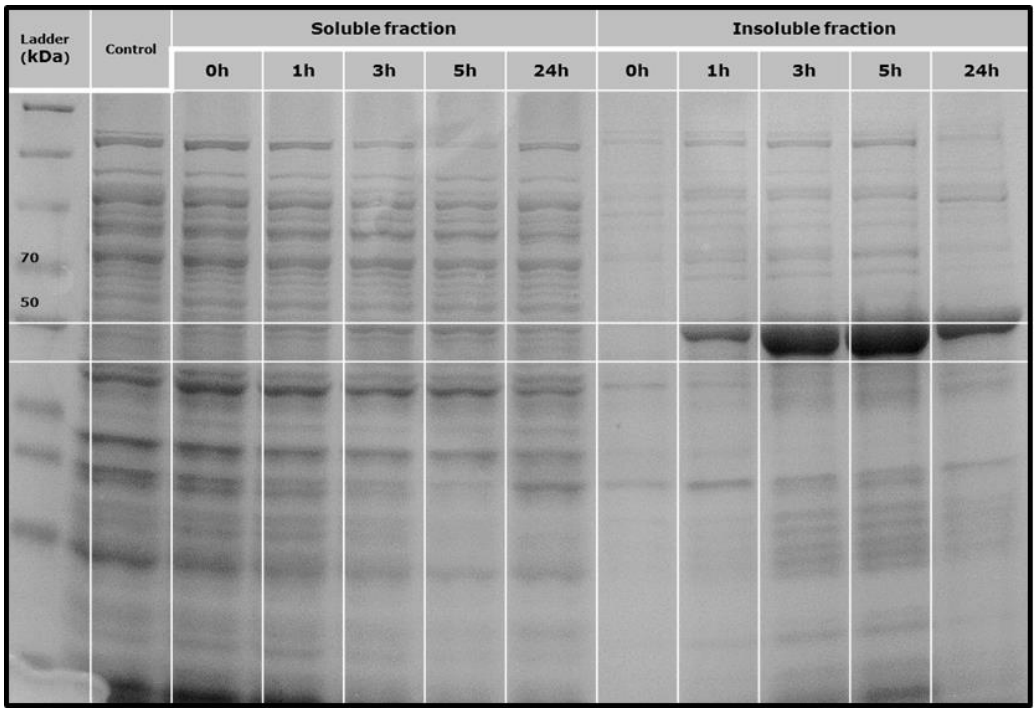


Figure 9.5: Expression of *A. niger* OhbA1
Expression of the enzyme at 30°C (A) and 37°C (B).

A


Ladder (kDa)
Control
Soluble fraction
Insoluble fraction
0h
1h
3h
5h
24h
0h
1h
3h
5h
24h
70
50

B


Ladder (kDa)
Control
Soluble fraction
Insoluble fraction
0h
1h
3h
5h
24h
0h
1h
3h
5h
24h
70
50

C

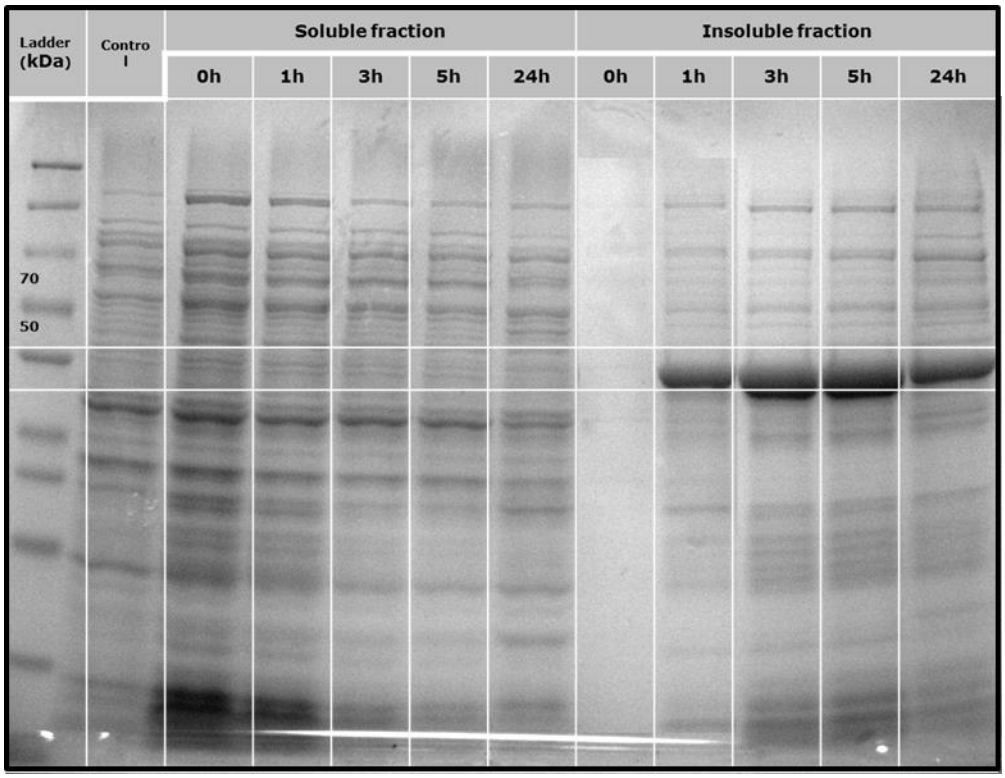


Figure 9.6: Expression of *A. terreus* CadA1
Expression of the enzyme at 20°C (A), 30°C (B) and 37°C (C).

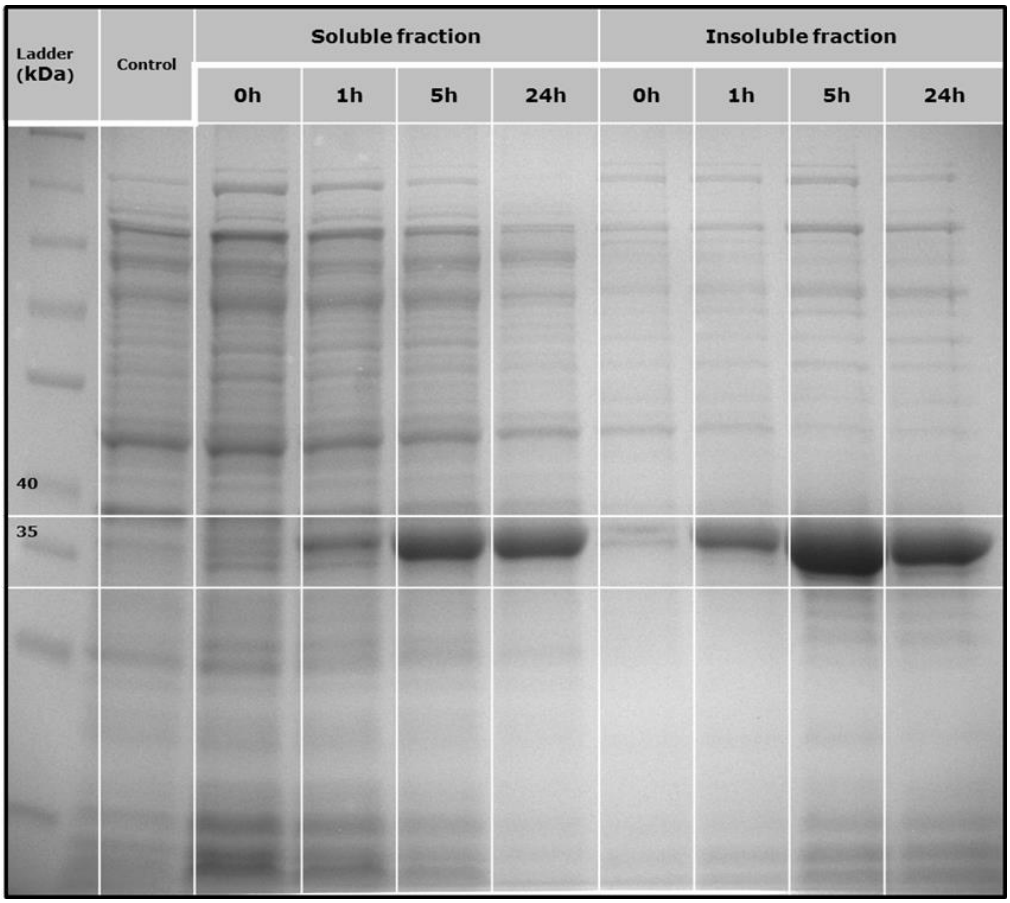
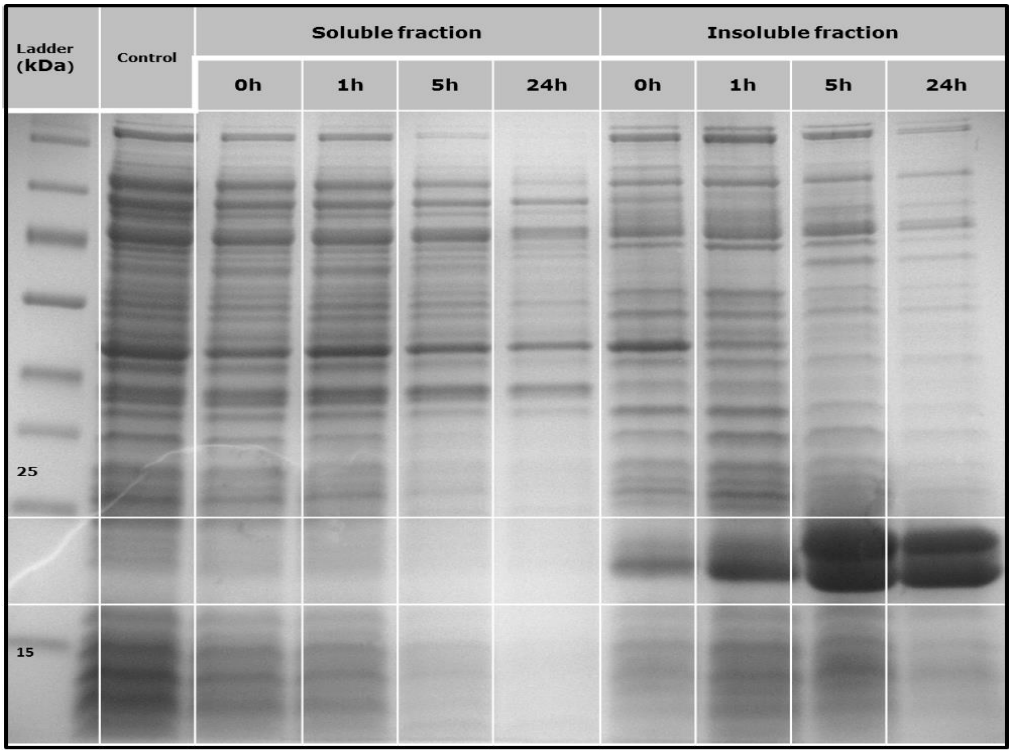


Figure 9.7: Expression of *P. fluorescens* ACMSD
Expression of the enzyme at 28˚C.

A



B

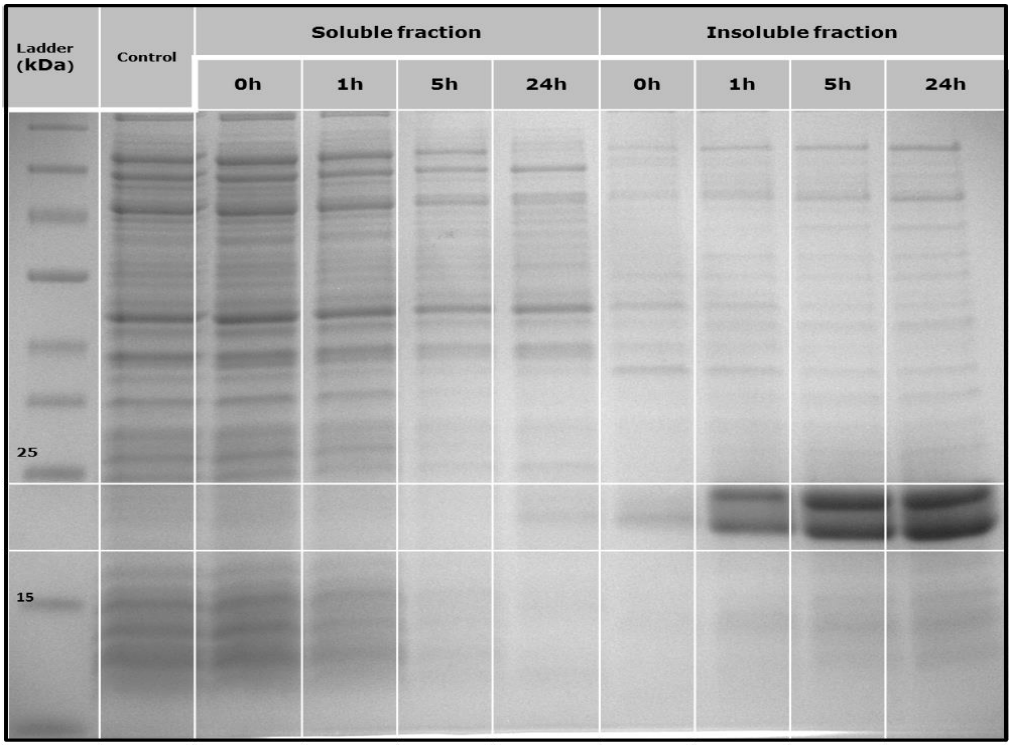
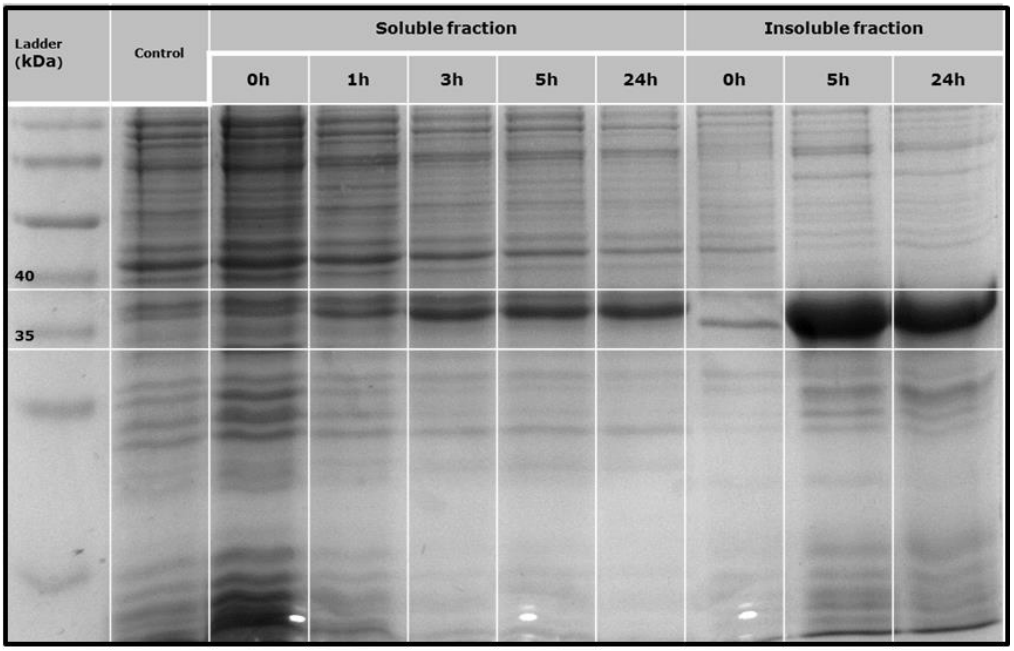


Figure 9.8: Expression of *P. fluorescens* HAO
Expression of the enzyme at 28˚C (A) and 18˚C (B).

A



B

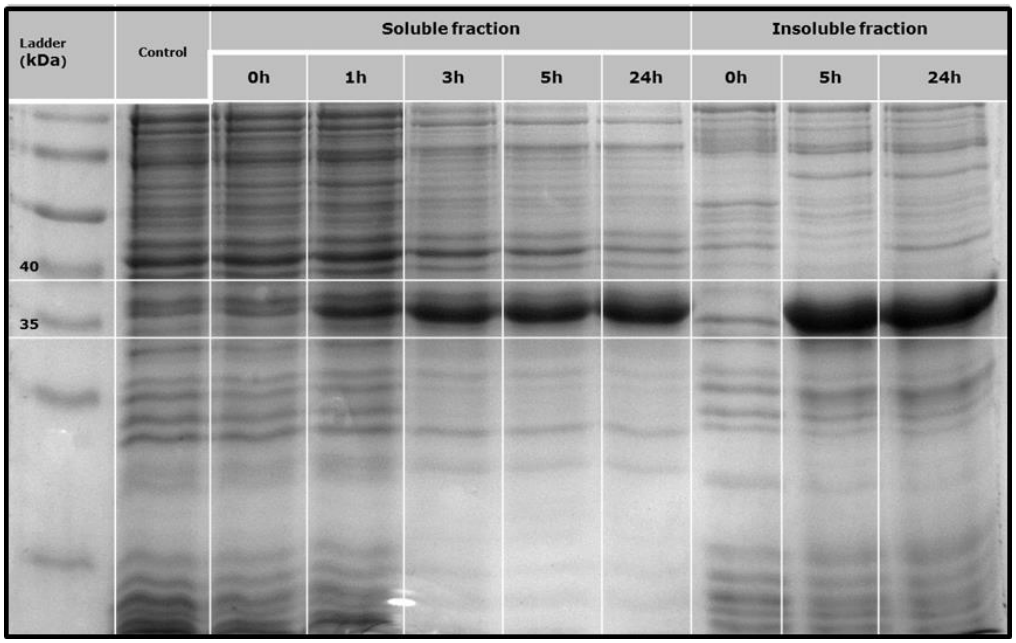


Figure 9.9: Expression of *P. torridus* MVD
Expression of the enzyme at 30°C (A) and 37°C (B).

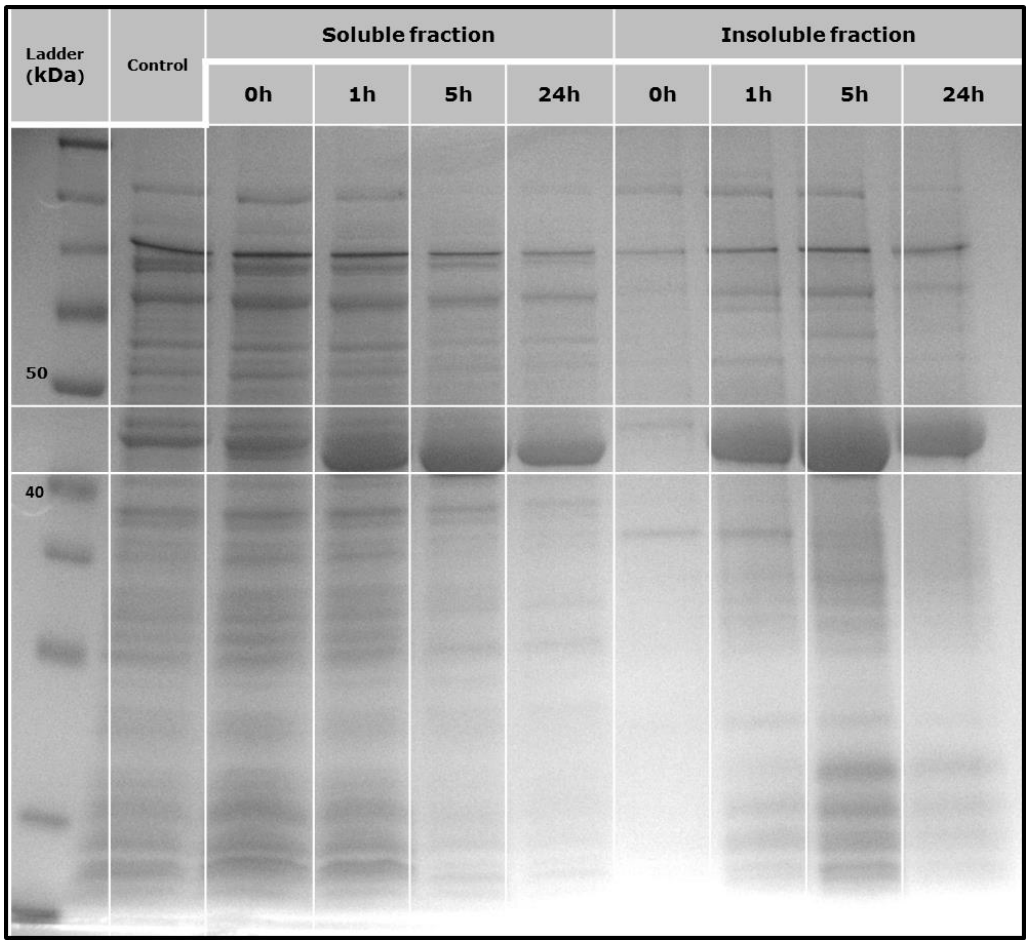


Figure 9.10: Expression of *S. cerevisiae* MVD
Expression of the enzyme at 37˚C.

9.4 Calibration curves

Standard samples for the compounds of interest were always analysed together with test samples every time an assay was performed. A calibration curve was calculated and used for quantification. By doing this, calibration curves were new every time and quantification never relied on dated data. Below are typical calibration curves for all compounds of interest.

For HPLC data (methacrylic, itaconic, mesaconic, citraconic, citramalic, sorbic and *cis*-aconitic acid quantification), values are indicated as an average of three replicates.

For GC-MS data (styrene, 1,3-pentadiene and isobutene quantification), values are indicated as an average of two replicates. For 1,3-pentadiene and styrene, samples were always run in the presence of internal standard (*tert*-butyl benzene) to minimize the error. The values indicated in the graphs are normalized to the internal standard.

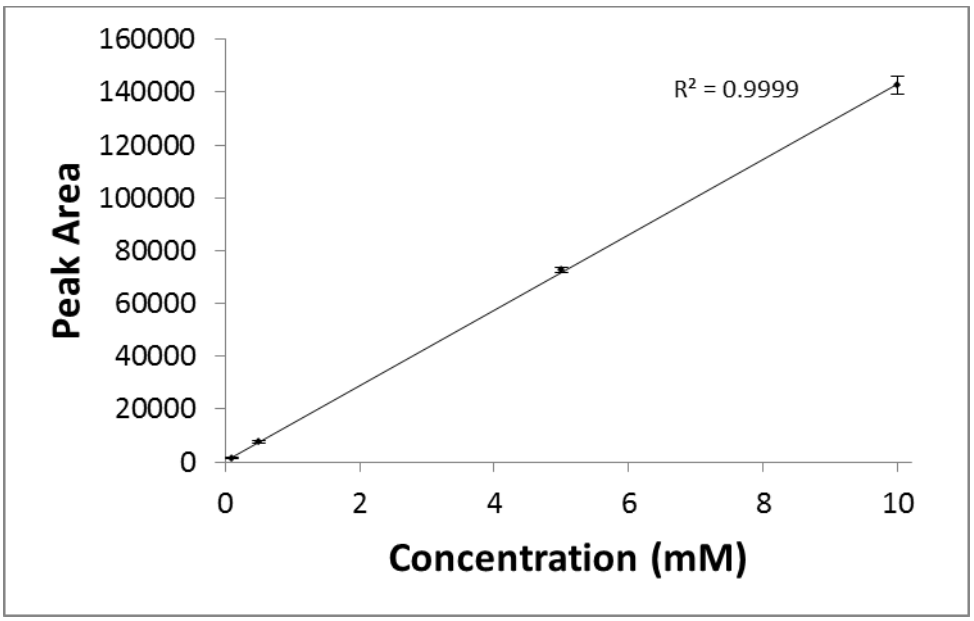


Figure 9.11: Methacrylic acid calibration curve

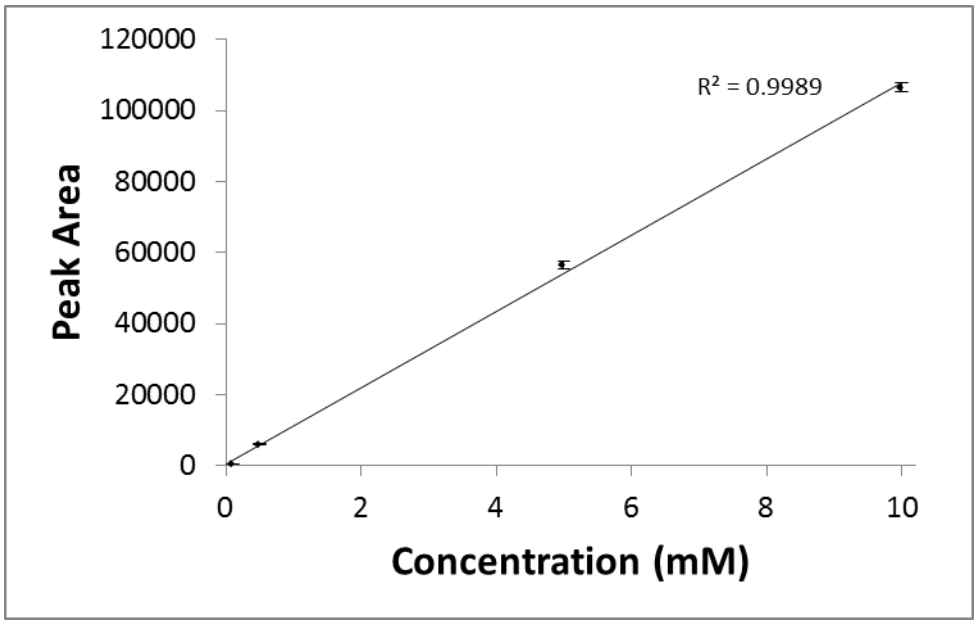


Figure 9.12: Itaconic acid calibration curve

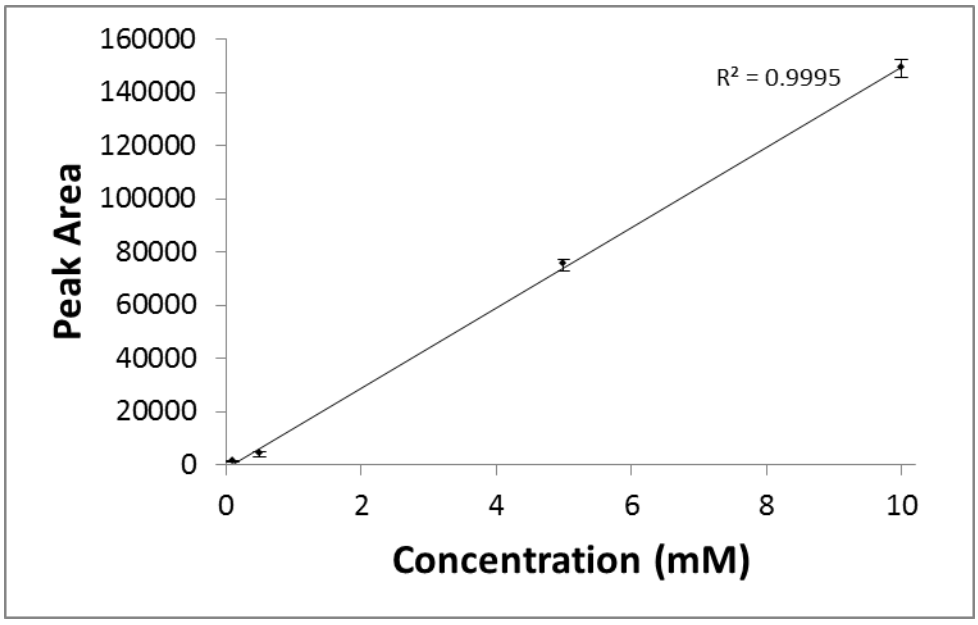


Figure 9.13: Mesaconic acid calibration curve

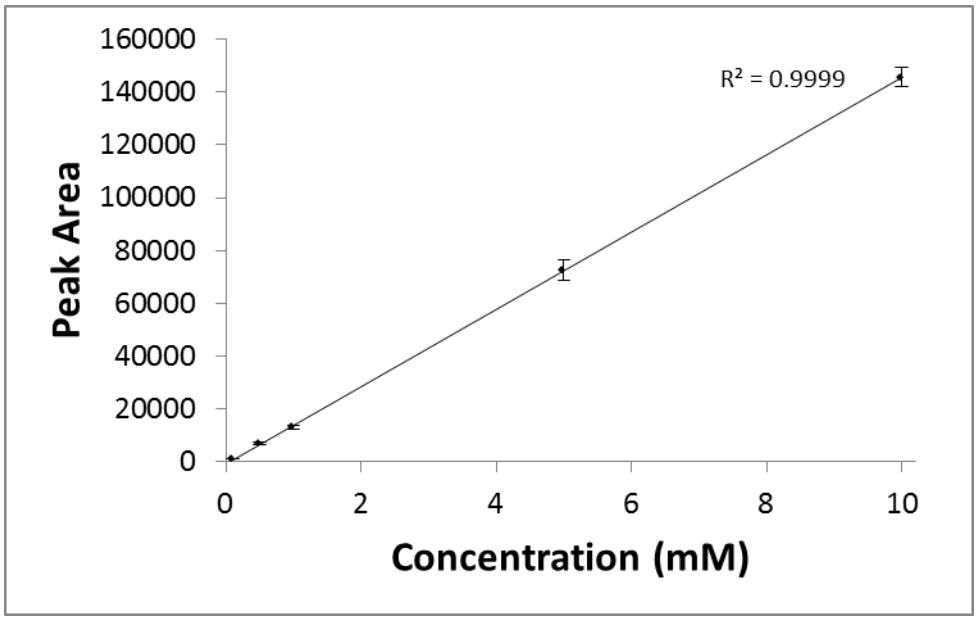


Figure 9.14: Citraconic acid calibration curve

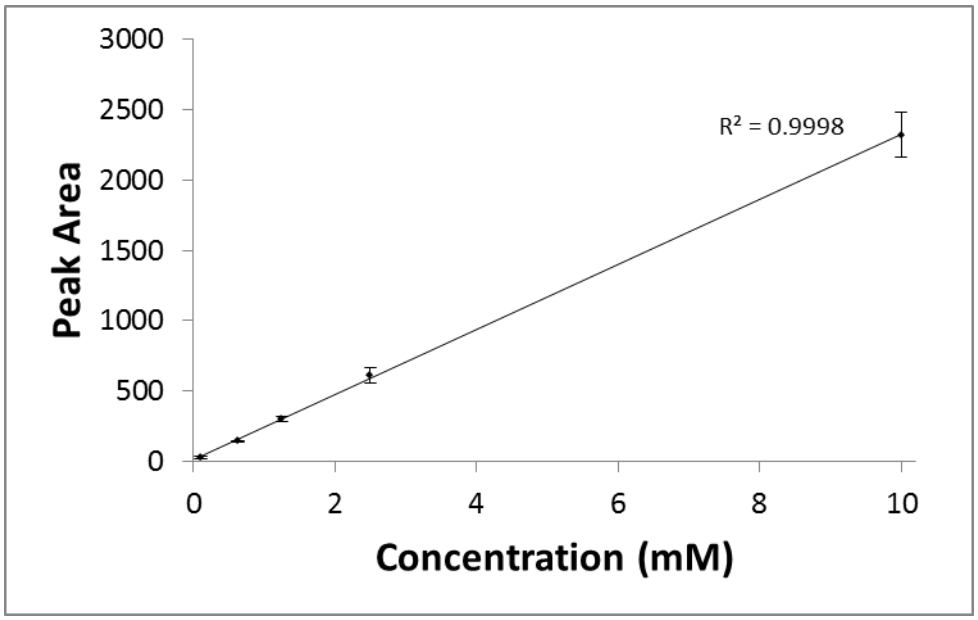


Figure 9.15: Citramalic acid calibration curve

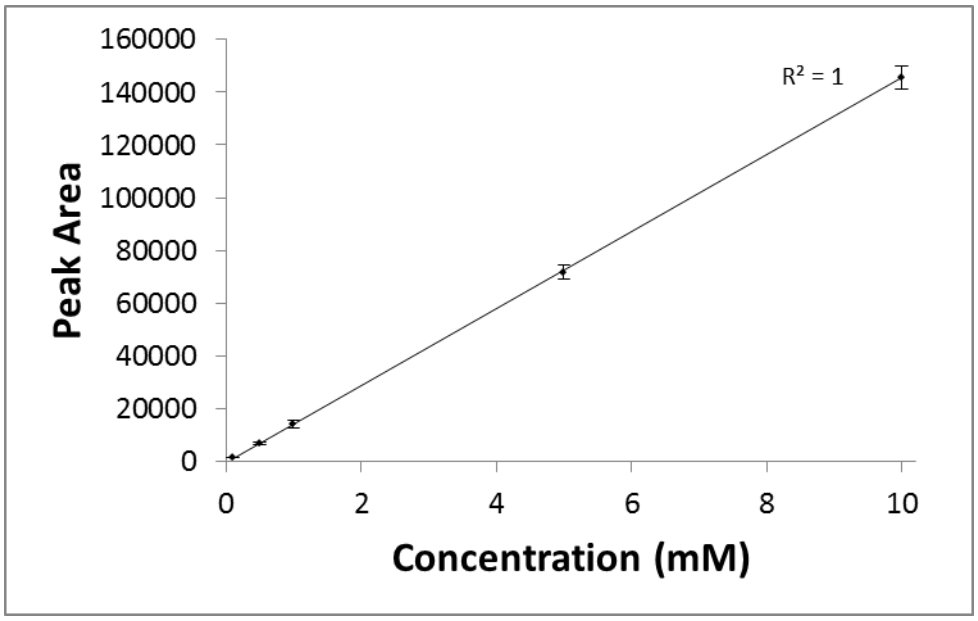


Figure 9.16: *cis*-Aconitic acid calibration curve

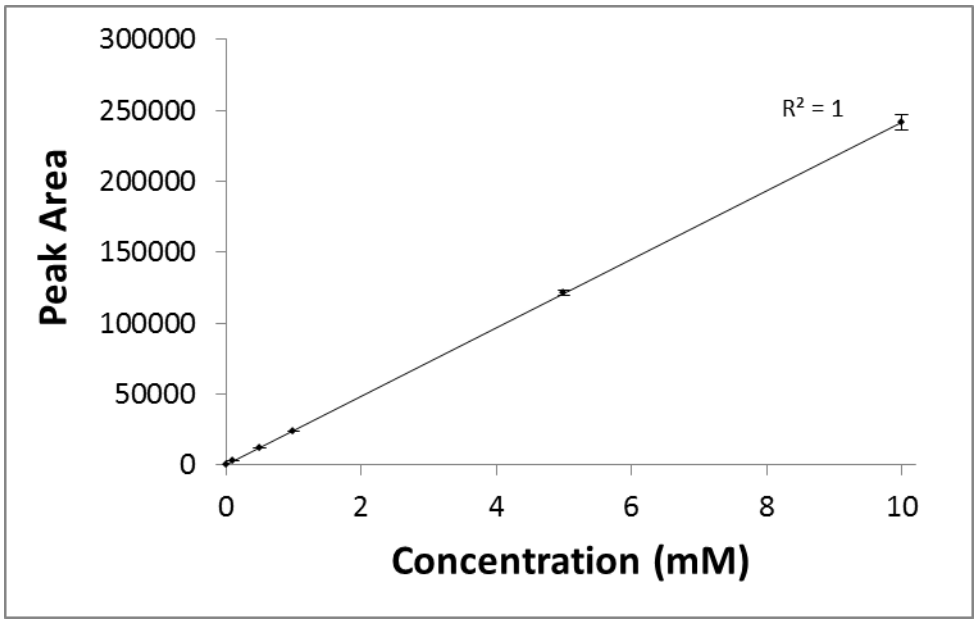


Figure 9.17: Fumaric acid calibration curve

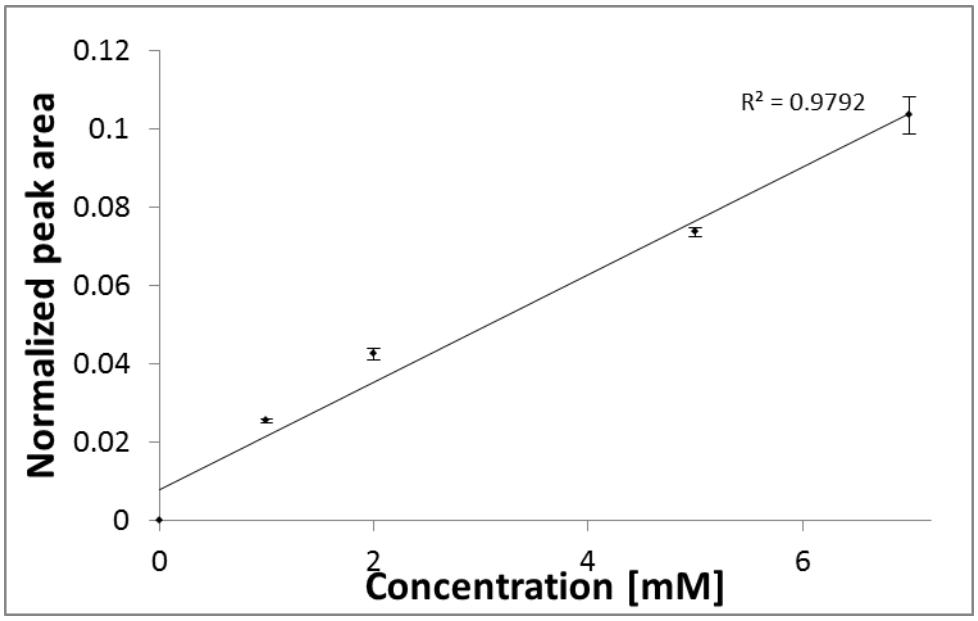


Figure 9.18: Styrene calibration curve

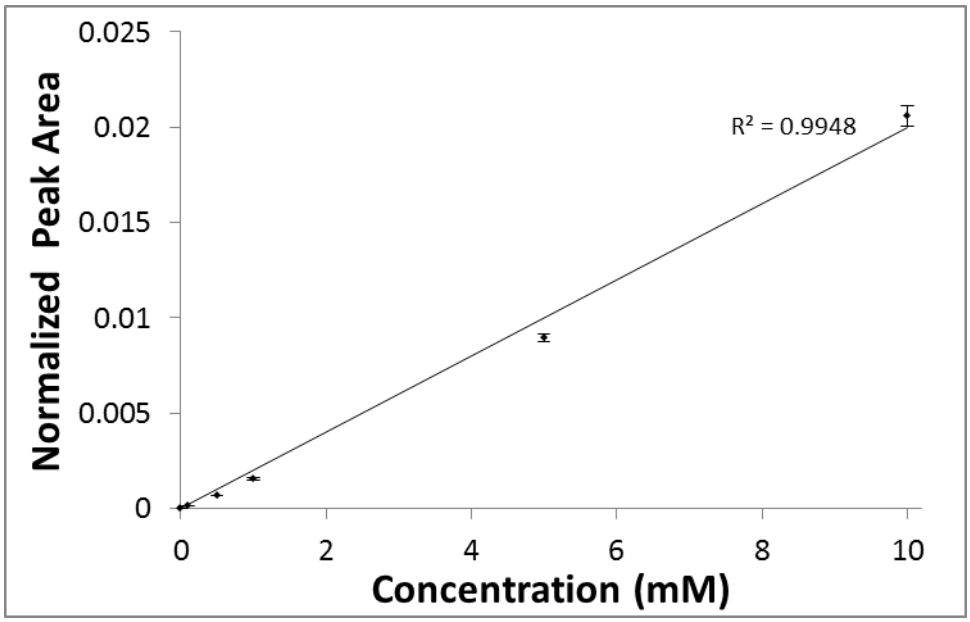


Figure 9.19: 1,3-Pentadiene calibration curve

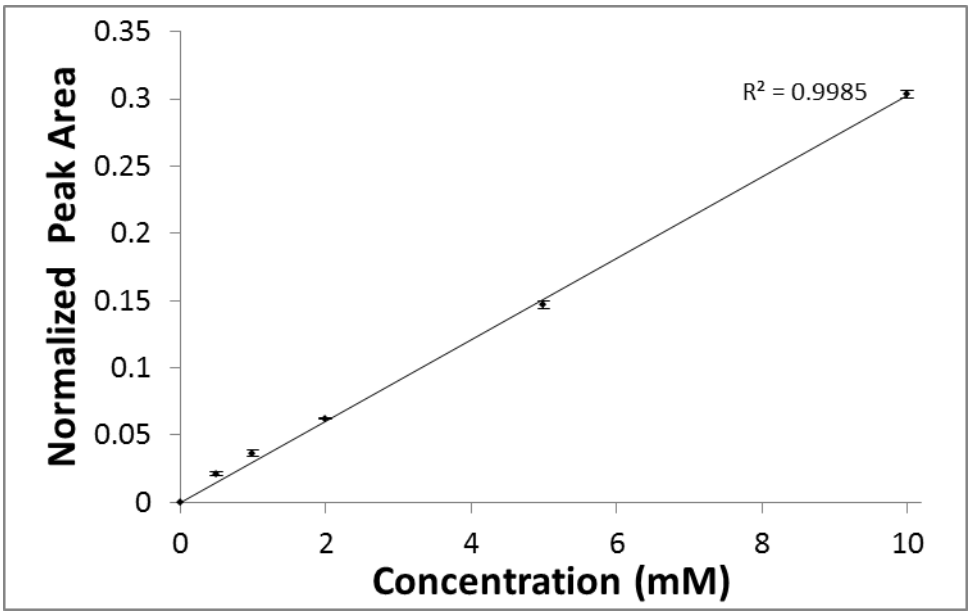


Figure 9.20: Isoprenol calibration curve

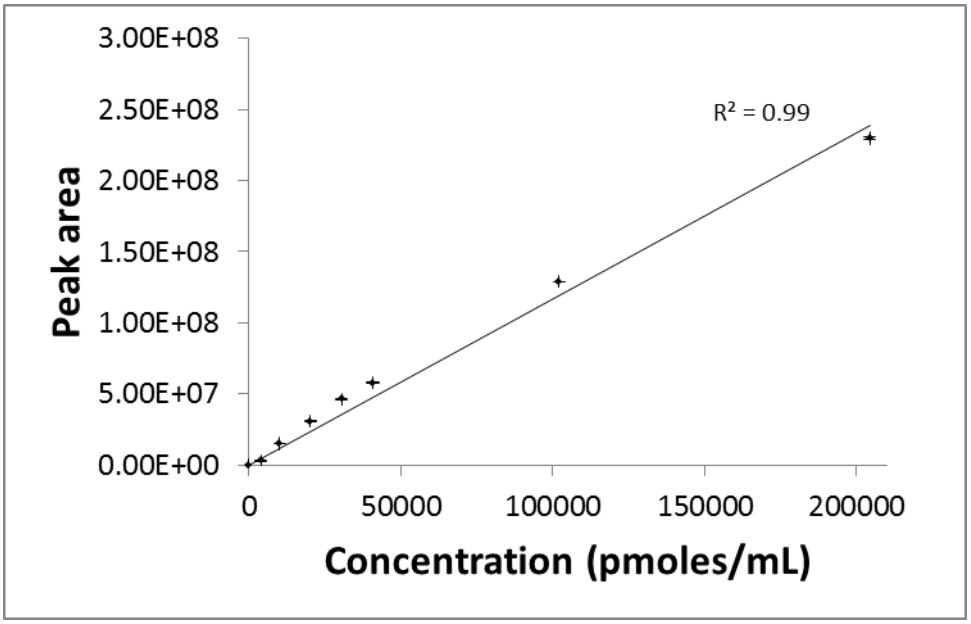


Figure 9.21: Isobutene calibration curve

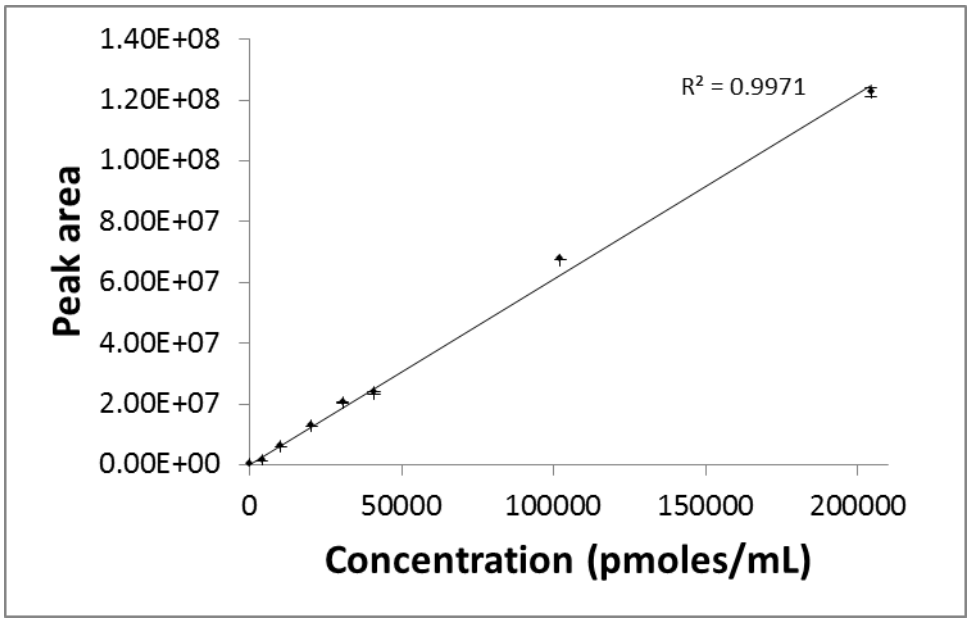


Figure 9.22: Propene calibration curve

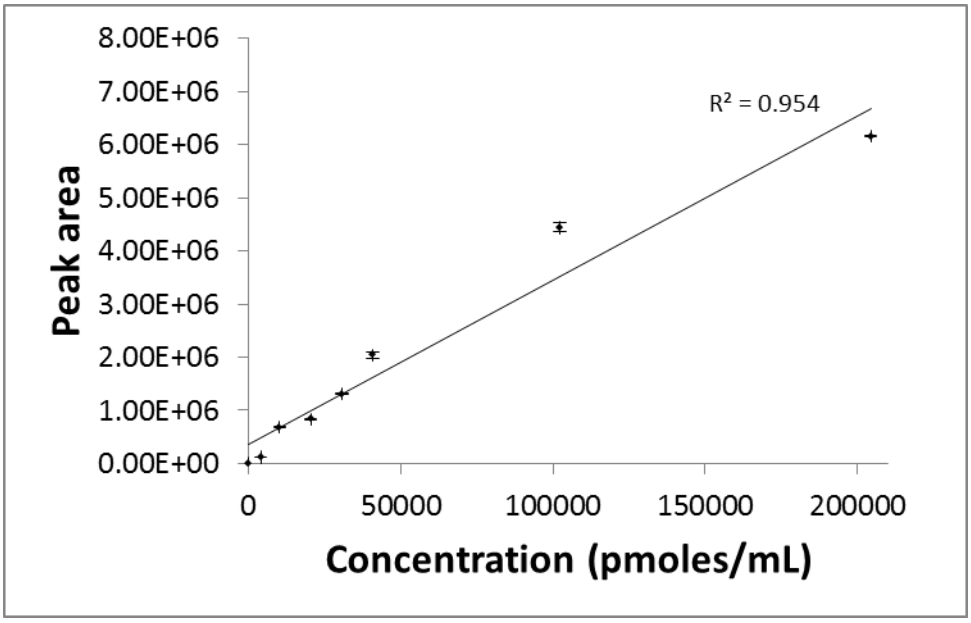


Figure 9.23: Ethene calibration curve

9.5 Raw data of Microresp™ screening

Table 9.1: *P. torridus* MVD variants (Library A)

Plate	1	2	3	4	5	6	7	8	9	10	11	12
Cut Off	0.692	0.790	0.647	0.690	0.801	0.774	0.786	0.713	0.646	0.652	0.646	0.670
DI	0.799	0.831	0.793	0.797	0.988	0.837	0.838	0.82	0.805	0.411	0.804	0.823
	0.794	0.814	0.8	0.76	0.871	0.828	0.81	0.841	0.809	0.762	0.802	0.805
	0.809	0.851	0.812	0.794	0.962	0.864	0.858	0.857	0.815	0.792	0.83	0.826
	0.817	0.86	0.812	0.791	0.955	0.86	0.895	0.861	0.846	0.347	0.842	0.823
	0.799	0.846	0.794	0.775	0.884	0.858	0.511	0.864	0.811	0.784	0.804	0.818
	0.788	0.806	0.797	0.762	0.815	0.838	0.802	0.809	0.794	0.798	0.774	0.809
	0.801	0.86	0.808	0.792	0.955	0.874	0.863	0.888	0.825	0.816	0.814	0.819
	0.816	0.982	0.818	0.811	0.976	0.876	0.88	0.901	0.836	0.816	0.83	0.822
	0.808	0.88	0.806	0.813	0.915	0.882	0.843	0.877	0.834	0.825	0.84	0.841
	0.786	0.832	0.797	0.779	0.823	0.813	0.772	0.838	0.78	0.795	0.77	0.798
	0.784	0.891	0.803	0.801	0.924	0.897	0.858	0.829	0.794	0.785	0.799	0.853
	0.793	0.817	0.808	0.772	0.972	0.875	0.435	0.836	0.822	0.79	0.812	0.837
	0.809	0.857	0.813	0.767	0.929	0.924	0.853	0.851	0.807	0.79	0.815	0.833
	0.812	0.856	0.791	0.777	0.993	0.864	0.861	0.837	0.823	0.775	0.821	0.845
	0.799	0.847	0.809	0.797	0.929	0.886	0.842	0.885	0.815	0.814	0.801	0.826
	0.794	0.872	0.805	0.69	0.89	0.86	0.879	0.865	0.811	0.817	0.33	0.833
	0.801	0.882	0.806	0.774	0.94	0.881	0.878	0.869	0.827	0.812	0.812	0.834
	0.805	0.947	0.826	0.816	0.96	0.882	0.875	0.909	0.842	0.839	0.838	0.841
	0.806	0.914	0.815	0.811	0.982	0.922	0.822	0.877	0.834	0.814	0.841	0.835
	0.822	0.934	0.821	0.837	0.952	0.88	0.339	0.86	0.826	0.819	0.835	0.83
	0.8	0.896	0.812	0.832	0.963	0.876	0.827	0.861	0.806	0.833	0.807	0.835
	0.807	0.806	0.805	0.746	0.891	0.844	0.345	0.849	0.817	0.808	0.832	0.844
	0.807	0.862	0.804	0.778	0.977	0.893	0.838	0.85	0.826	0.801	0.834	0.842
	0.817	0.883	0.813	0.809	0.852	0.861	0.889	0.856	0.813	0.81	0.83	0.845
	0.796	0.852	0.811	0.79	0.911	0.896	0.832	0.892	0.798	0.819	0.81	0.845
	0.489	0.868	0.825	0.797	0.952	0.856	0.857	0.877	0.827	0.828	0.818	0.849
	0.818	0.86	0.816	0.817	0.967	0.864	0.863	0.886	0.83	0.822	0.832	0.841
	0.809	0.956	0.826	0.575	0.912	0.877	0.827	0.888	0.817	0.819	0.835	0.85
	0.827	0.95	0.806	0.821	0.887	0.861	0.337	0.904	0.838	0.826	0.848	0.866
	0.825	0.941	0.827	0.869	0.963	0.9	0.818	0.875	0.859	0.817	0.874	0.846
	0.799	0.882	0.804	0.762	0.974	0.923	0.871	0.82	0.824	0.803	0.847	0.842
	0.804	0.888	0.812	0.289	0.869	0.873	0.882	0.846	0.805	0.752	0.812	0.835
	0.799	0.857	0.808	0.754	0.981	0.883	0.881	0.87	0.834	0.813	0.874	0.797
	0.811	0.854	0.806	0.805	0.981	0.889	0.872	0.854	0.835	0.783	0.828	0.832
	0.805	0.874	0.815	0.8	0.933	0.856	0.855	0.858	0.804	0.801	0.812	0.823
	0.803	0.878	0.824	0.86	0.933	0.853	0.851	0.875	0.817	0.852	0.829	0.643
	0.82	0.888	0.821	0.811	0.93	0.838	0.852	0.902	0.824	0.818	0.823	0.845
	0.815	0.999	0.832	0.856	0.933	0.851	0.88	0.902	0.831	0.811	0.839	0.864
	0.805	0.921	0.828	0.824	0.912	0.86	0.84	0.893	0.828	0.354	0.853	0.833
	0.819	0.942	0.841	0.904	0.999	0.859	0.878	0.893	0.85	0.835	0.863	0.847
	0.812	0.919	0.809	0.818	0.993	0.894	0.913	0.911	0.839	0.79	0.855	0.842
	0.8	0.878	0.83	0.789	0.983	0.94	0.909	0.92	0.855	0.334	0.866	0.859
	0.812	0.884	0.821	0.792	0.922	0.869	0.915	0.888	0.833	0.775	0.857	0.835
	0.824	0.888	0.821	0.795	0.92	0.84	0.911	0.802	0.825	0.803	0.818	0.85
	0.812	0.882	0.816	0.844	0.887	0.892	0.871	0.847	0.832	0.74	0.829	0.827
	0.808	0.928	0.827	0.843	0.923	0.855	0.886	0.871	0.819	0.819	0.799	0.849
	0.827	0.889	0.838	0.857	0.976	0.833	0.85	0.879	0.836	0.822	0.835	0.831
	0.831	0.954	0.837	0.85	0.974	0.886	0.848	0.874	0.846	0.83	0.833	0.861
	0.84	0.972	0.847	0.344	0.975	0.89	0.873	0.909	0.827	0.82	0.818	0.84
	0.829	0.845	0.854	0.872	0.869	0.861	0.929	0.891	0.842	0.838	0.84	0.856
	0.805	0.926	0.82	0.87	0.942	0.883	0.935	0.933	0.846	0.831	0.838	0.855
	0.816	0.859	0.831	0.813	0.87	0.901	0.906	0.91	0.828	0.769	0.823	0.839
	0.814	0.921	0.84	0.831	0.874	0.859	0.971	0.936	0.861	0.83	0.871	0.854
	0.834	0.922	0.83	0.311	0.999	0.859	0.924	0.889	0.846	0.36	0.83	0.849
	0.819	0.892	0.835	0.814	0.934	0.791	0.876	0.924	0.828	0.813	0.847	0.842
	0.807	0.917	0.824	0.831	0.925	0.907	0.861	0.876	0.846	0.822	0.823	0.847
	0.847	0.929	0.828	0.835	0.921	0.992	0.986	0.875	0.837	0.826	0.826	0.919
	0.796	0.678	0.835	0.844	0.888	0.896	0.911	0.902	0.853	0.824	0.846	0.868
	0.832	0.979	0.834	0.833	0.887	0.932	0.885	0.913	0.839	0.839	0.825	0.885
	0.839	0.944	0.848	0.853	0.886	0.988	0.853	0.9	0.862	0.823	0.803	0.869
	0.818	0.947	0.825	0.795	0.952	0.875	0.97	0.877	0.847	0.847	0.377	0.866
	0.821	0.925	0.851	0.856	0.874	0.879	0.964	0.945	0.897	0.84	0.87	0.851
	0.825	0.947	0.839	0.841	0.963	0.841	0.894	0.934	0.893	0.808	0.89	0.864
	0.843	0.96	0.824	0.829	0.9	0.969	0.993	0.911	0.863	0.801	0.844	0.87
	0.832	0.953	0.843	0.832	0.933	0.968	0.962	0.877	0.862	0.834	0.82	0.851
	0.824	0.913	0.821	0.833	0.933	0.962	0.461	0.911	0.851	0.834	0.384	0.858
	0.83	0.942	0.83	0.839	0.93	0.956	0.995	0.914	0.842	0.823	0.824	0.865
	0.84	0.874	0.846	0.855	0.933	0.955	0.955	0.928	0.843	0.835	0.84	0.906
	0.851	0.921	0.846	0.87	0.912	0.95	0.944	0.937	0.854	0.815	0.815	0.91
	0.86	0.999	0.846	0.862	0.999	0.841	0.949	0.903	0.852	0.366	0.353	0.886
	0.827	0.916	0.836	0.369	0.993	0.856	0.908	0.853	0.827	0.852	0.787	0.843
	0.822	0.93	0.85	0.846	0.983	0.921	0.992	0.926	0.849	0.819	0.864	0.859
	0.833	0.95	0.841	0.846	0.922	0.601	0.988	0.927	0.869	0.381	0.895	0.863
	0.86	0.875	0.837	0.857	0.883	0.995	0.892	0.914	0.835	0.815	0.838	0.879
	0.824	0.892	0.85	0.88	0.901	0.955	0.955	0.939	0.844	0.855	0.367	0.856
	0.831	0.956	0.84	0.856	0.859	0.944	0.947	0.917	0.85	0.849	0.825	0.85
	0.84	0.964	0.837	0.871	0.859	0.949	0.936	0.931	0.832	0.832	0.843	0.882
	0.869	0.912	0.842	0.91	0.791	0.908	0.922	0.929	0.843	0.854	0.869	0.867
	0.863	0.968	0.855	0.898	0.907	0.992	0.978	0.914	0.871	0.858	0.86	0.861
	0.885	0.932	0.845	0.917	0.992	0.988	0.933	0.964	0.864	0.815	0.822	0.844

Table 9.2: *P. torridus* MVD variants (Library B)

Plate	1	2	3	4	5	6	7	8	9	10	11	12
Cut Off	0.740	0.702	0.683	0.789	0.640	0.792	0.656	0.635	0.706	0.683	0.739	0.787
DI	0.848	0.774	0.774	0.839	0.788	0.839	0.773	0.791	0.83	0.803	0.816	0.878
	0.845	0.766	0.765	0.839	0.686	0.889	0.795	0.783	0.805	0.807	0.845	0.872
	0.884	0.788	0.815	0.872	0.704	0.908	0.822	0.803	0.837	0.837	0.827	0.929
	0.933	0.796	0.818	0.911	0.656	0.918	0.824	0.812	0.848	0.845	0.882	0.958
	0.9	0.785	0.805	0.868	0.684	0.909	0.822	0.792	0.827	0.826	0.858	0.93
	0.832	0.746	0.794	0.816	0.728	0.821	0.8	0.789	0.833	0.785	0.824	0.857
	0.943	0.793	0.832	0.887	0.719	0.924	0.833	0.811	0.849	0.822	0.87	0.931
	0.905	0.778	0.869	0.877	0.852	0.892	0.85	0.816	0.865	0.869	0.903	0.883
	0.881	0.802	0.84	0.896	0.885	0.877	0.842	0.823	0.892	0.854	0.918	0.933
	0.823	0.758	0.8	0.804	0.753	0.799	0.77	0.776	0.798	0.798	0.837	0.847
	0.925	0.746	0.809	0.83	0.797	0.866	0.355	0.762	0.843	0.834	0.865	0.946
	0.588	0.748	0.781	0.89	0.751	0.901	0.812	0.771	0.824	0.836	0.839	0.873
	0.906	0.782	0.807	0.881	0.756	0.887	0.809	0.778	0.823	0.798	0.83	0.922
	0.909	0.784	0.81	0.882	0.749	0.915	0.831	0.797	0.825	0.82	0.85	0.902
	0.895	0.779	0.841	0.895	0.77	0.95	0.862	0.8	0.823	0.868	0.89	0.898
	0.922	0.761	0.834	0.545	0.774	0.881	0.84	0.814	0.863	0.844	0.865	0.867
	0.946	0.744	0.833	0.89	0.836	0.923	0.835	0.801	0.827	0.802	0.873	0.906
	0.913	0.777	0.871	0.887	0.835	0.858	0.871	0.837	0.904	0.854	0.898	0.874
	0.924	0.79	0.852	0.919	0.88	0.879	0.846	0.814	0.888	0.834	0.878	0.884
	0.909	0.802	0.861	0.884	0.883	0.873	0.83	0.815	0.856	0.861	0.902	0.89
	0.926	0.74	0.827	0.836	0.44	0.878	0.412	0.768	0.82	0.855	0.878	0.914
	0.818	0.74	0.773	0.818	0.72	0.858	0.803	0.746	0.828	0.816	0.816	0.838
	0.893	0.782	0.803	0.949	0.741	0.958	0.8	0.774	0.799	0.783	0.853	0.909
	0.895	0.795	0.815	0.896	0.733	0.901	0.841	0.501	0.827	0.823	0.877	0.917
	0.88	0.768	0.824	0.871	0.769	0.896	0.828	0.785	0.807	0.833	0.832	0.834
	0.906	0.777	0.854	0.888	0.804	0.902	0.864	0.801	0.784	0.847	0.881	0.885
	0.917	0.799	0.849	0.905	0.809	0.897	0.845	0.813	0.845	0.802	0.895	0.927
	0.889	0.745	0.882	0.868	0.887	0.896	0.865	0.83	0.939	0.879	0.887	0.885
	0.905	0.794	0.873	0.881	0.91	0.886	0.87	0.82	0.898	0.883	0.926	0.927
	0.915	0.741	0.887	0.926	0.841	0.944	0.788	0.822	0.899	0.925	0.856	0.927
	0.875	0.756	0.816	0.851	0.822	0.938	0.821	0.772	0.819	0.831	0.853	0.895
	0.924	0.729	0.807	0.812	0.721	0.913	0.803	0.788	0.8	0.802	0.843	0.936
	0.422	0.771	0.86	0.913	0.781	0.941	0.828	0.78	0.805	0.816	0.861	0.876
	0.898	0.754	0.838	0.915	0.777	0.936	0.84	0.77	0.797	0.834	0.882	0.876
	0.902	0.755	0.841	0.501	0.719	0.883	0.834	0.783	0.827	0.824	0.848	0.9
	0.936	0.733	0.884	0.867	0.787	0.853	0.852	0.784	0.859	0.829	0.921	0.905
	0.923	0.757	0.867	0.846	0.903	0.842	0.854	0.813	0.873	0.85	0.883	0.903
	0.909	0.768	0.873	0.849	0.886	0.884	0.883	0.802	0.877	0.869	0.913	0.885
	0.919	0.757	0.89	0.878	0.955	0.402	0.885	0.79	0.868	0.885	0.909	0.873
	0.919	0.601	0.89	0.862	0.926	0.882	0.813	0.798	0.885	0.901	0.907	0.872
	0.897	0.764	0.841	0.878	0.757	0.904	0.818	0.773	0.834	0.827	0.871	0.946
	0.863	0.398	0.84	0.938	0.772	0.998	0.857	0.774	0.837	0.868	0.9	0.924
	0.866	0.755	0.826	0.873	0.775	0.897	0.827	0.799	0.794	0.823	0.874	0.882
	0.941	0.775	0.845	0.859	0.74	0.876	0.863	0.795	0.809	0.85	0.895	0.951
	0.919	0.766	0.854	0.889	0.765	0.872	0.859	0.784	0.796	0.857	0.892	0.918
	0.88	0.742	0.861	0.851	0.728	0.828	0.871	0.806	0.881	0.86	0.893	0.902
	0.893	0.763	0.852	0.861	0.843	0.84	0.854	0.819	0.85	0.867	0.872	0.876
	0.898	0.728	0.861	0.898	0.847	0.941	0.856	0.813	0.865	0.893	0.901	0.878
	0.927	0.803	0.887	0.892	0.854	0.912	0.906	0.839	0.894	0.904	0.937	0.924
	0.934	0.81	0.888	0.868	0.854	0.897	0.875	0.839	0.91	0.909	0.908	0.905
	0.917	0.783	0.86	0.872	0.945	0.912	0.87	0.817	0.837	0.853	0.902	0.954
	0.857	0.522	0.832	0.488	0.856	0.896	0.867	0.796	0.838	0.878	0.92	0.889
	0.922	0.819	0.869	0.874	0.863	0.401	0.885	0.818	0.827	0.901	0.957	0.921
	0.943	0.745	0.846	0.873	0.76	0.916	0.847	0.789	0.825	0.588	0.885	0.995
	0.888	0.79	0.854	0.879	0.86	0.937	0.889	0.819	0.863	0.885	0.908	0.9
	0.912	0.774	0.827	0.884	0.815	0.922	0.853	0.822	0.848	0.855	0.883	0.919
	0.923	0.789	0.841	0.901	0.809	0.988	0.851	0.853	0.898	0.862	0.88	0.963
	0.962	0.801	0.88	0.898	0.908	0.991	0.949	0.828	0.879	0.502	0.927	0.947
	0.995	0.819	0.865	0.903	0.89	0.942	0.879	0.861	0.923	0.886	0.911	0.976
	0.997	0.833	0.888	0.947	0.929	0.956	0.881	0.815	0.858	0.853	0.903	0.985
	0.957	0.779	0.829	0.863	0.096	0.921	0.845	0.841	0.873	0.841	0.906	0.902
	0.957	0.787	0.862	0.902	0.763	0.912	0.878	0.826	0.85	0.906	0.962	0.876
	0.993	0.829	0.862	0.886	0.842	0.902	0.87	0.846	0.814	0.871	0.967	0.878
	0.995	0.789	0.88	0.975	0.776	0.978	0.866	0.844	0.837	0.824	0.915	0.924
	0.921	0.816	0.844	0.964	0.782	0.996	0.859	0.856	0.892	0.85	0.86	0.905
	0.969	0.838	0.814	0.994	0.784	0.501	0.87	0.876	0.907	0.872	0.894	0.982
	0.976	0.823	0.859	0.965	0.842	0.918	0.88	0.863	0.928	0.859	0.514	0.87
	0.957	0.855	0.918	0.975	0.9	0.911	0.911	0.873	0.514	0.969	0.958	0.866
	0.993	0.866	0.911	0.964	0.913	0.898	0.887	0.881	0.962	0.762	0.972	0.859
	0.995	0.946	0.898	0.994	0.918	0.863	0.885	0.855	0.938	0.909	0.959	0.87
	0.922	0.818	0.863	0.965	0.784	0.857	0.871	0.775	0.846	0.848	0.921	0.88
	0.969	0.821	0.857	0.966	0.842	0.841	0.903	0.826	0.866	0.926	0.971	0.947
	0.98	0.818	0.841	0.975	0.9	0.798	0.875	0.827	0.876	0.839	0.922	0.863
	0.957	0.736	0.798	0.964	0.807	0.898	0.877	0.847	0.884	0.839	0.918	0.902
	0.993	0.821	0.894	0.994	0.91	0.927	0.903	0.852	0.889	0.945	0.964	0.886
	0.943	0.837	0.88	0.965	0.905	0.934	0.877	0.868	0.892	0.92	0.945	0.975
	0.888	0.844	0.918	0.913	0.826	0.917	0.882	0.866	0.927	0.946	0.963	0.964
	0.912	0.836	0.999	0.918	0.827	0.857	0.933	0.92	0.932	0.947	0.878	0.994
	0.923	0.884	0.94	0.822	0.847	0.922	0.898	0.883	0.943	0.955	0.992	0.355
	0.952	0.873	0.398	0.867	0.852	0.922	0.888	0.895	0.968	0.993	0.846	0.987

Table 9.3: *S. cerevisiae* MVD variants (Library A)

Plate	1	2	3	4	5	6	7	8	9	10	11	12
Cut Off	0.607	0.645	0.636	0.646	0.637	0.638	0.685	0.666	0.616	0.619	0.612	0.672
DI	0.837	0.794	0.827	0.823	0.828	0.83	0.841	0.881	0.813	0.82	0.818	0.828
	0.808	0.808	0.808	0.806	0.443	0.81	0.813	0.873	0.795	0.802	0.789	0.807
	0.791	0.804	0.817	0.831	0.825	0.807	0.831	0.871	0.808	0.815	0.806	0.839
	0.792	0.817	0.827	0.844	0.834	0.595	0.83	0.865	0.808	0.821	0.821	0.82
	0.783	0.818	0.832	0.822	0.837	0.573	0.835	0.869	0.804	0.812	0.822	0.455
	0.336	0.804	0.83	0.836	0.802	0.822	0.82	0.853	0.799	0.805	0.798	0.818
	0.781	0.818	0.83	0.869	0.826	0.824	0.814	0.854	0.785	0.812	0.805	0.82
	0.782	0.82	0.834	0.828	0.834	0.824	0.818	0.861	0.808	0.828	0.792	0.824
	0.795	0.804	0.369	0.821	0.831	0.814	0.815	0.865	0.8	0.825	0.81	0.823
	0.788	0.304	0.819	0.805	0.811	0.827	0.834	0.847	0.793	0.795	0.794	0.817
	0.829	0.816	0.812	0.811	0.821	0.827	0.826	0.865	0.811	0.811	0.799	0.825
	0.812	0.802	0.822	0.831	0.836	0.83	0.815	0.859	0.805	0.812	0.803	0.815
	0.807	0.795	0.831	0.823	0.819	0.813	0.806	0.871	0.808	0.814	0.803	0.837
	0.8	0.81	0.833	0.833	0.831	0.815	0.822	0.854	0.81	0.809	0.803	0.828
	0.789	0.792	0.836	0.814	0.831	0.798	0.815	0.879	0.79	0.801	0.83	0.81
	0.784	0.805	0.834	0.822	0.821	0.814	0.834	0.862	0.796	0.396	0.828	0.814
	0.781	0.795	0.821	0.853	0.821	0.815	0.813	0.87	0.789	0.808	0.8	0.806
	0.797	0.803	0.829	0.827	0.826	0.826	0.813	0.848	0.814	0.813	0.804	0.816
	0.794	0.798	0.824	0.828	0.829	0.807	0.811	0.877	0.799	0.805	0.792	0.815
	0.793	0.8	0.818	0.824	0.818	0.817	0.377	0.86	0.792	0.814	0.797	0.807
	0.851	0.813	0.837	0.82	0.822	0.813	0.836	0.871	0.813	0.825	0.799	0.825
	0.811	0.794	0.849	0.841	0.829	0.82	0.822	0.865	0.801	0.815	0.802	0.37
	0.801	0.811	0.834	0.804	0.512	0.808	0.835	0.851	0.8	0.824	0.784	0.826
	0.805	0.832	0.845	0.828	0.833	0.806	0.818	0.858	0.806	0.827	0.819	0.831
	0.803	0.795	0.835	0.805	0.823	0.496	0.828	0.865	0.807	0.822	0.816	0.819
	0.798	0.802	0.836	0.827	0.828	0.844	0.814	0.609	0.506	0.806	0.805	0.821
	0.79	0.383	0.821	0.82	0.823	0.816	0.818	0.861	0.81	0.8	0.351	0.803
	0.775	0.801	0.816	0.827	0.815	0.821	0.815	0.845	0.802	0.815	0.826	0.825
	0.796	0.8	0.831	0.828	0.818	0.824	0.8	0.839	0.808	0.823	0.817	0.812
	0.775	0.838	0.836	0.84	0.821	0.8	0.824	0.866	0.805	0.817	0.794	0.846
	0.821	0.805	0.82	0.816	0.814	0.805	0.818	0.867	0.824	0.353	0.797	0.818
	0.809	0.801	0.811	0.84	0.821	0.826	0.812	0.85	0.8	0.797	0.801	0.814
	0.821	0.823	0.838	0.83	0.839	0.792	0.818	0.551	0.811	0.82	0.346	0.825
	0.822	0.824	0.827	0.819	0.822	0.811	0.823	0.86	0.804	0.82	0.813	0.809
	0.804	0.796	0.825	0.815	0.795	0.822	0.83	0.879	0.824	0.811	0.818	0.8
	0.793	0.826	0.823	0.814	0.815	0.82	0.825	0.881	0.811	0.79	0.818	0.804
	0.808	0.794	0.806	0.861	0.81	0.823	0.432	0.868	0.805	0.791	0.795	0.793
	0.781	0.812	0.82	0.788	0.815	0.821	0.824	0.846	0.809	0.81	0.804	0.815
	0.801	0.818	0.825	0.814	0.793	0.807	0.827	0.844	0.804	0.801	0.809	0.797
	0.805	0.826	0.821	0.834	0.797	0.816	0.832	0.848	0.813	0.806	0.814	0.651
	0.86	0.809	0.841	0.82	0.834	0.811	0.833	0.854	0.817	0.825	0.802	0.83
	0.808	0.807	0.832	0.84	0.831	0.799	0.837	0.876	0.809	0.807	0.795	0.822
	0.793	0.82	0.816	0.839	0.839	0.383	0.824	0.858	0.82	0.808	0.41	0.83
	0.794	0.822	0.82	0.82	0.826	0.803	0.825	0.879	0.423	0.802	0.822	0.83
	0.781	0.807	0.82	0.818	0.825	0.817	0.827	0.882	0.808	0.806	0.827	0.821
	0.805	0.815	0.791	0.823	0.834	0.811	0.823	0.859	0.807	0.8	0.833	0.817
	0.786	0.294	0.803	0.841	0.822	0.815	0.815	0.847	0.797	0.803	0.81	0.806
	0.8	0.82	0.822	0.844	0.546	0.479	0.422	0.858	0.808	0.816	0.806	0.808
	0.789	0.818	0.818	0.808	0.829	0.824	0.823	0.86	0.815	0.803	0.788	0.401
	0.767	0.827	0.83	0.847	0.825	0.834	0.836	0.851	0.815	0.815	0.805	0.797
	0.842	0.823	0.834	0.828	0.841	0.823	0.827	0.87	0.803	0.81	0.809	0.82
	0.84	0.819	0.847	0.839	0.84	0.819	0.836	0.876	0.815	0.828	0.809	0.811
	0.809	0.814	0.837	0.864	0.843	0.8	0.838	0.868	0.828	0.823	0.355	0.846
	0.792	0.823	0.815	0.835	0.826	0.834	0.656	0.855	0.806	0.81	0.802	0.836
	0.808	0.802	0.794	0.823	0.823	0.828	0.834	0.58	0.811	0.807	0.809	0.821
	0.811	0.836	0.368	0.822	0.832	0.823	0.811	0.848	0.818	0.806	0.808	0.819
	0.803	0.815	0.788	0.851	0.612	0.826	0.714	0.848	0.812	0.805	0.803	0.802
	0.811	0.826	0.391	0.829	0.826	0.838	0.833	0.867	0.829	0.822	0.846	0.817
	0.813	0.833	0.825	0.822	0.844	0.833	0.821	0.842	0.799	0.325	0.341	0.814
	0.804	0.862	0.825	0.814	0.834	0.858	0.843	0.842	0.814	0.816	0.784	0.824
	0.817	0.834	0.844	0.851	0.834	0.824	0.842	0.864	0.829	0.818	0.834	0.822
	0.819	0.82	0.84	0.843	0.846	0.786	0.84	0.872	0.807	0.819	0.836	0.832
	0.822	0.392	0.841	0.836	0.841	0.817	0.855	0.861	0.831	0.849	0.805	0.834
	0.798	0.814	0.828	0.829	0.843	0.819	0.838	0.871	0.825	0.824	0.831	0.426
	0.834	0.816	0.819	0.837	0.838	0.835	0.588	0.862	0.826	0.814	0.856	0.474
	0.817	0.827	0.807	0.828	0.831	0.85	0.843	0.863	0.841	0.835	0.818	0.84
	0.808	0.823	0.818	0.863	0.822	0.838	0.82	0.877	0.821	0.813	0.806	0.824
	0.82	0.829	0.82	0.829	0.842	0.841	0.851	0.864	0.839	0.844	0.838	0.844
	0.82	0.841	0.826	0.839	0.846	0.838	0.828	0.867	0.347	0.814	0.812	0.84
	0.818	0.866	0.838	0.838	0.832	0.848	0.85	0.854	0.826	0.404	0.814	0.822
	0.816	0.814	0.836	0.841	0.84	0.836	0.84	0.886	0.821	0.825	0.82	0.84
	0.811	0.831	0.834	0.833	0.836	0.82	0.608	0.88	0.81	0.81	0.548	0.814
	0.811	0.828	0.844	0.839	0.836	0.845	0.839	0.884	0.824	0.841	0.828	0.853
	0.788	0.824	0.84	0.849	0.833	0.844	0.834	0.876	0.825	0.828	0.858	0.853
	0.796	0.828	0.844	0.837	0.818	0.848	0.839	0.879	0.808	0.832	0.855	0.835
	0.811	0.842	0.831	0.847	0.84	0.818	0.836	0.88	0.825	0.828	0.846	0.857
	0.817	0.835	0.809	0.865	0.834	0.836	0.836	0.885	0.819	0.829	0.817	0.85
	0.804	0.846	0.834	0.843	0.846	0.842	0.851	0.901	0.83	0.84	0.833	0.851
	0.812	0.833	0.832	0.835	0.846	0.841	0.836	0.889	0.811	0.851	0.809	0.846
	0.821	0.839	0.841	0.839	0.835	0.851	0.856	0.896	0.814	0.85	0.82	0.857

Table 9.4: *S. cerevisiae* MVD variants (Library B)

Plate	1	2	3	4	5	6	7	8	9	10	11	12
Cut Off	0.655	0.631	0.640	0.606	0.678	0.635	0.659	0.621	0.623	0.631	0.656	0.638
DI	0.819	0.823	0.486	0.819	0.869	0.835	0.833	0.815	0.813	0.831	0.829	0.866
	0.809	0.809	0.809	0.789	0.861	0.818	0.825	0.8	0.811	0.801	0.81	0.579
	0.815	0.831	0.831	0.795	0.884	0.827	0.846	0.824	0.81	0.813	0.832	0.83
	0.82	0.831	0.828	0.803	0.87	0.841	0.852	0.82	0.814	0.802	0.833	0.549
	0.813	0.83	0.819	0.807	0.857	0.834	0.836	0.799	0.811	0.801	0.833	0.853
	0.809	0.812	0.372	0.799	0.852	0.406	0.836	0.796	0.814	0.797	0.821	0.823
	0.806	0.822	0.434	0.791	0.868	0.828	0.85	0.791	0.809	0.842	0.835	0.51
	0.821	0.827	0.823	0.806	0.867	0.833	0.849	0.824	0.818	0.807	0.842	0.547
	0.828	0.823	0.814	0.806	0.867	0.825	0.827	0.811	0.821	0.815	0.852	0.503
	0.812	0.808	0.551	0.801	0.861	0.813	0.843	0.801	0.802	0.805	0.83	0.814
	0.832	0.391	0.826	0.796	0.85	0.821	0.819	0.811	0.803	0.811	0.813	0.844
	0.816	0.836	0.829	0.777	0.86	0.822	0.818	0.805	0.793	0.793	0.805	0.828
	0.815	0.825	0.817	0.791	0.872	0.817	0.831	0.816	0.796	0.812	0.834	0.842
	0.82	0.83	0.824	0.8	0.867	0.509	0.843	0.82	0.809	0.799	0.838	0.859
	0.81	0.827	0.819	0.809	0.869	0.821	0.835	0.809	0.805	0.798	0.853	0.86
	0.816	0.812	0.825	0.803	0.866	0.831	0.845	0.813	0.805	0.796	0.83	0.828
	0.809	0.829	0.822	0.791	0.851	0.545	0.837	0.799	0.802	0.816	0.83	0.836
	0.818	0.377	0.488	0.795	0.865	0.377	0.847	0.811	0.83	0.829	0.846	0.846
	0.823	0.48	0.81	0.792	0.864	0.817	0.842	0.806	0.821	0.811	0.849	0.434
	0.828	0.464	0.814	0.79	0.865	0.819	0.845	0.803	0.808	0.824	0.828	0.806
	0.826	0.563	0.832	0.792	0.857	0.378	0.847	0.803	0.801	0.83	0.818	0.846
	0.819	0.831	0.823	0.794	0.874	0.821	0.849	0.826	0.81	0.811	0.837	0.497
	0.818	0.833	0.812	0.797	0.869	0.833	0.831	0.81	0.811	0.814	0.838	0.835
	0.826	0.833	0.627	0.794	0.881	0.828	0.85	0.816	0.807	0.792	0.848	0.84
	0.827	0.562	0.808	0.807	0.828	0.821	0.813	0.81	0.816	0.797	0.835	0.835
	0.826	0.82	0.806	0.794	0.87	0.822	0.85	0.818	0.817	0.792	0.862	0.838
	0.823	0.674	0.809	0.778	0.87	0.815	0.837	0.816	0.806	0.829	0.848	0.815
	0.818	0.43	0.819	0.798	0.868	0.832	0.839	0.811	0.79	0.837	0.833	0.83
	0.83	0.832	0.835	0.796	0.87	0.835	0.866	0.813	0.799	0.82	0.843	0.838
	0.823	0.828	0.83	0.796	0.859	0.836	0.838	0.808	0.807	0.824	0.834	0.823
	0.816	0.814	0.817	0.796	0.864	0.829	0.835	0.816	0.802	0.828	0.817	0.833
	0.824	0.812	0.816	0.797	0.851	0.821	0.829	0.811	0.808	0.807	0.815	0.81
	0.816	0.787	0.826	0.8	0.849	0.833	0.847	0.805	0.797	0.81	0.822	0.859
	0.831	0.815	0.819	0.806	0.865	0.837	0.853	0.814	0.813	0.797	0.834	0.85
	0.827	0.815	0.398	0.814	0.875	0.817	0.836	0.811	0.832	0.806	0.834	0.843
	0.824	0.807	0.833	0.794	0.861	0.836	0.863	0.818	0.837	0.796	0.844	0.846
	0.817	0.803	0.817	0.796	0.88	0.827	0.828	0.797	0.8	0.807	0.839	0.494
	0.837	0.818	0.83	0.798	0.873	0.833	0.847	0.821	0.803	0.828	0.831	0.465
	0.836	0.834	0.825	0.79	0.877	0.836	0.85	0.809	0.81	0.798	0.832	0.84
	0.827	0.821	0.831	0.792	0.869	0.843	0.854	0.803	0.811	0.823	0.83	0.82
	0.825	0.822	0.835	0.791	0.87	0.855	0.84	0.815	0.81	0.83	0.828	0.852
	0.823	0.824	0.497	0.802	0.861	0.833	0.839	0.817	0.795	0.8	0.832	0.83
	0.875	0.815	0.819	0.806	0.866	0.844	0.846	0.82	0.798	0.81	0.813	0.856
	0.825	0.821	0.822	0.813	0.873	0.835	0.855	0.825	0.819	0.793	0.82	0.84
	0.832	0.816	0.476	0.807	0.876	0.833	0.837	0.814	0.814	0.808	0.844	0.843
	0.843	0.798	0.378	0.798	0.892	0.818	0.856	0.818	0.841	0.798	0.836	0.839
	0.81	0.803	0.823	0.794	0.881	0.823	0.833	0.804	0.809	0.816	0.831	0.5
	0.837	0.831	0.828	0.818	0.872	0.842	0.841	0.817	0.831	0.827	0.845	0.838
	0.833	0.834	0.819	0.795	0.862	0.389	0.84	0.816	0.817	0.801	0.832	0.83
	0.822	0.827	0.785	0.801	0.867	0.848	0.849	0.802	0.828	0.824	0.843	0.843
	0.816	0.826	0.845	0.81	0.859	0.839	0.866	0.821	0.809	0.799	0.833	0.645
	0.833	0.836	0.836	0.8	0.888	0.845	0.838	0.837	0.819	0.807	0.862	0.854
	0.833	0.846	0.839	0.811	0.886	0.842	0.864	0.838	0.804	0.81	0.866	0.853
	0.823	0.82	0.831	0.8	0.889	0.833	0.849	0.823	0.802	0.803	0.825	0.844
	0.842	0.82	0.836	0.807	0.879	0.833	0.849	0.819	0.803	0.812	0.832	0.845
	0.846	0.819	0.827	0.814	0.878	0.838	0.86	0.821	0.809	0.808	0.862	0.853
	0.846	0.819	0.823	0.796	0.887	0.822	0.82	0.818	0.81	0.855	0.823	0.418
	0.837	0.828	0.843	0.812	0.881	0.515	0.852	0.846	0.809	0.87	0.848	0.838
	0.852	0.834	0.831	0.824	0.876	0.844	0.855	0.833	0.815	0.828	0.858	0.836
	0.843	0.807	0.503	0.815	0.866	0.829	0.85	0.819	0.838	0.817	0.856	0.846
	0.837	0.855	0.856	0.815	0.883	0.853	0.848	0.825	0.807	0.818	0.834	0.875
	0.83	0.826	0.833	0.802	0.871	0.845	0.845	0.814	0.809	0.806	0.828	0.873
	0.837	0.831	0.831	0.805	0.874	0.845	0.843	0.832	0.825	0.809	0.841	0.871
	0.83	0.827	0.383	0.818	0.893	0.853	0.864	0.83	0.831	0.811	0.842	0.872
	0.839	0.828	0.833	0.822	0.886	0.842	0.84	0.844	0.818	0.82	0.833	0.862
	0.849	0.834	0.76	0.81	0.893	0.827	0.851	0.839	0.82	0.801	0.854	0.854
	0.826	0.511	0.373	0.807	0.877	0.413	0.85	0.808	0.821	0.833	0.841	0.84
	0.847	0.583	0.836	0.81	0.89	0.828	0.863	0.834	0.829	0.845	0.844	0.852
	0.855	0.821	0.847	0.813	0.888	0.832	0.875	0.842	0.831	0.821	0.851	0.862
	0.839	0.452	0.848	0.833	0.887	0.464	0.855	0.816	0.83	0.83	0.859	0.845
	0.846	0.859	0.848	0.802	0.882	0.847	0.833	0.82	0.823	0.81	0.831	0.862
	0.828	0.84	0.858	0.804	0.874	0.842	0.833	0.833	0.82	0.828	0.825	0.875
	0.834	0.509	0.85	0.821	0.893	0.843	0.846	0.855	0.838	0.842	0.838	0.702
	0.824	0.404	0.842	0.811	0.899	0.831	0.857	0.829	0.817	0.861	0.841	0.868
	0.823	0.833	0.848	0.826	0.88	0.83	0.862	0.82	0.826	0.823	0.863	0.871
	0.84	0.601	0.843	0.818	0.888	0.844	0.862	0.832	0.818	0.823	0.857	0.449
	0.819	0.827	0.387	0.818	0.891	0.842	0.859	0.82	0.867	0.851	0.85	0.538
	0.841	0.848	0.826	0.827	0.903	0.852	0.875	0.836	0.84	0.851	0.854	0.439
	0.841	0.829	0.84	0.822	0.903	0.862	0.864	0.823	0.846	0.815	0.864	0.862
	0.842	0.82	0.843	0.822	0.891	0.872	0.854	0.816	0.842	0.844	0.869	0.856

9.6 Search for homologues

Table 9.5: Homologues of *P. torridus* MVD [a].

	Organism	NCBI Reference	Identity (%)	Similarity (%)	E value
1	***Picrophilus torridus***	**YP_024134**	**100**	**100**	**-**
2	***Ferroplasma acidarmanus***	**YP_008142696**	**58.02**	**76.23**	**2.00E-131**
3	***Thermoplasma volcanium***	**NP_110839**	**42.86**	**66.46**	**1.00E-83**
4	***Thermoplasma acidophilum***	**NP_394760**	**38.73**	**63.81**	**8.00E-70**
5	***Thermoplasmatales archaeon A-plasma***	**WP_021795407**	**53.7**	**75.93**	**3.00E-11**
6	*Lactobacillus versmoldensis*	WP_010623536	22.47	43.99	8.00E-06
7	*Lactobacillus iners*	WP_006736906	22.29	42.52	3.00E-05
8	*Streptococcus entericus*	WP_018367392	22.92	40.86	4.00E-04
9	*Encephalitozoon intestinalis*	XP_003073009	24.76	40.84	5.00E-04
10	*Candidatus Micrarchaeum acidiphilum*	EET89942	23.2	43.89	9.00E-04
11	*Halonotius sp.*	WP_021040093	23.24	41.9	0.001
12	*Salinarchaeum sp.*	YP_008055205	19.57	43.48	0.002
13	*Enterococcus faecium*	WP_002351456	19.03	41.69	0.002
14	*Streptococcus parauberis*	WP_003104536	25.08	42.67	0.002
15	*Marinilabilia salmonicolor*	WP_010661914	24.63	44.83	0.002
16	*Borrelia afzelii*	YP_006931874	23.55	43.43	0.002
17	*Helcococcus kunzii*	WP_005398566	22.57	45.14	0.002
18	*Borrelia garinii*	YP_073129	25	45.67	0.003
19	*Waddlia chondrophila*	YP_003709481	25.47	47.2	0.003
20	*Haloferax elongans*	WP_008324111	22.26	40.85	0.003
21	*Natronomonas moolapensis*	YP_007486922	21.54	40	0.004
22	*Haloferax gibbonsii*	WP_004977100	21.54	41.85	0.006
23	*Gardnerella vaginalis*	YP_003373567	24.86	46.49	0.006
24	*Lactobacillus plantarum*	YP_007986982	20.8	40.67	0.008
25	*Halorubrum lipolyticum*	WP_008006049	24.38	39.67	0.009
26	*Halorhabdus utahensi*	YP_003129610	21.88	39.38	0.01
27	*Borrelia bissettii*	YP_004777918	30.19	47.17	0.012
28	*Methanomassiliicoccus luminyensis*	WP_019177897	21.89	42.42	0.016
29	*Halococcus thailandensis*	WP_007742266	22.39	40.8	0.016
30	*Pneumocystis jirovecii*	CCJ31570	22.71	43	0.017
31	*Haloferax prahovense*	WP_008093241	21.81	41.43	0.018
32	*Halogranum salarium*	WP_009374450	22.56	40.85	0.021
33	*Acholeplasma palmae*	CCV64523	21.45	44.16	0.022
34	*Guillardia theta*	XP_005819249	21.9	39.54	0.023
35	*Halococcus morrhuae*	WP_004055468	21.98	41.49	0.024
36	*Haloferax larsenii*	WP_007540006	21.6	41.05	0.028
37	*Borrelia burgdorferi*	YP_005806868	26.15	47.25	0.031
38	*Halococcus*	WP_007692401	21.47	40.49	0.032

	hamelinensis				
39	*Haloarcula vallismortis*	WP_004518165	22.26	40.24	0.035
40	*Haloferax volcanii*	YP_003535463	21.28	41.03	0.035
41	*Lactobacillus ruminis*	YP_004832120	22.9	44.84	0.036
42	*Lactobacillus vaginalis*	WP_003716957	25.41	42.7	0.05
43	*Streptococcus macacae*	WP_003081606	22.04	41.45	0.055
44	*Kazachstania africana*	XP_003956599	23.45	42.41	0.059
45	*Pneumocystis murina*	EMR10307	21.81	42.68	0.072
46	*Haloferax sulfurifontis*	WP_007273787	21.54	41.85	0.086
47	*Halobacterium salinarum*	YP_001688719	21.9	39.68	0.098

[a] *Thermoplasmatales* proteins are in bold

Table 9.6: Homologoues of *T. acidophilum* mevalonate-3-phosphate-5-kinase

Organism	NCBI Reference	Identity (%)	Similarity (%)	E value
Thermoplasma acidophilum	WP_010901177	100	100	-
Thermoplasma volcanium	BAB60008	54.92	74.61	5.00E-72
Thermoplasmatales archaeon A-plasma	WP_021794774	52.88	68.06	2.00E-64
Picrophilus torridus	**WP_011177280**	**55.49**	**72.56**	**2.00E-56**
Ferroplasma sp. Type II	WP_021787503	46.56	69.31	4.00E-54
Thermoplasmatales archaeon E-plasma	WP_021792357	45.79	66.32	9.00E-53
Ferroplasma acidarmanus	WP_009886959	45.5	68.78	1.00E-51

[a] The *P. torridus* homologue is characterized by a level of identity of 55.5% (in bold).

10 Publications

Published papers:

ROSSONI, L., HALL, S. J., EASTHAM, G., LICENCE, P. & STEPHENS, G. 2015. The putative mevalonate diphosphate decarboxylase from *Picrophilus torridus* is in reality a mevalonate-3-kinase with high potential for bioproduction of isobutene. *Appl Environ Microbiol,* 81**,** 2625-2634.

Papers in draft:

A screening assay for decarboxylases based on monitoring gaseous carbon dioxide

Inhibition analysis of α -amino-β-carboxymuconate-ε-semialdehyde decarboxylase reveals new crucial characteristics for substrate binding and reactivity