



Yue Sun, BSc MSc

**Biocatalysts improvements:
Protein Engineering of Photodecarboxylases and Discovery
of Baeyer-Villiger Monooxygenases**

DOCTORAL THESIS

To achieve the University Degree of
Doktorin der Naturwissenschaften (Dr.rer.nat)

submitted to

Graz University of Technology

Supervisor

Univ.-Prof. Dr.rer.nat. Robert Kourist
Institute of Molecular Biotechnology

Graz, June and 2021

AFFIDAVIT

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present doctoral thesis.

Date, Signature

Abstract of dissertation

The term “Biocatalysis” denotes the utilizations of enzymes or whole cells as catalysts to access a broad range of chemical transformations. Historical major (perceived or real) shortcomings of biocatalysis are referred to as enzymes (catalyst) are expensive, unstable and inefficient. However, biocatalysis is a very dynamic field of research and development. Many of the previous shortcomings of biocatalysis are already resolved due to the speculator advances in molecular biology and biotechnology. The core step for biocatalysis engineering is discovering and engineering better performances (e.g., high activity, desired substrate selectivity, high enantioselectivity and high stability) enzymes. The improvement of biocatalysts could be a contributing factor to more successful applications in research and industry. This dissertation set out to access biocatalysts improvement for two different enzymes, that is the fatty acid photodecarboxylase (FAP, EC 4.1.1.106) and Baeyer-Villiger monooxygenase (EC 1.14.13.x).

In chapter I, through reconstructing the possible ancestral decarboxylases from a set of 12 extant fatty acid photodecarboxylase (FAP) sequences, three ancestral photodecarboxylases were resurrected. One of the resurrected enzymes (ANC1) showed higher stability, solubility but lower catalytic activity than extant FAP. ANC1 exhibits more hydrophilic amino acids on the protein surface, which might improve the solubility of the enzyme. The higher stability and lower catalytic activity might be attributed to there are more rigid regions for ANC1.

Further improvement of stability of CvFAP by rational design and site-directed mutagenesis achieved a thermostable variant CvFAP_{S473G}. The residue E550 of CvFAP plays a crucial role in catalytic activity. The variant sANC1_{V366L} improved the specific activity than sANC1_{WT}. By introducing deep eutectic solvents (DES) as an alternative solvent system, the catalytic activity for FAPs could be maintained and improved.

In chapter II, with gene screening, 18 putative BVMOs encoding genes were identified from the genome of fermentative bacteria *Cupriavidus necator* H16, which comprising efficient NADH regeneration dehydrogenases naturally. One of the BVMOs enables to perform the BV oxidation and sulfoxidation reaction. By overexpressing BVMO in *C.necator* H16, whole cell-mediated BV-oxidation does not need external NADH. With fluorescence assay, this novel BVMO presented a relatively high melting temperature.

The aims of this dissertation is to improve biocatalysts' desired properties with protein discovery and engineering. Within this work, since the photoenzyme was seldom found in nature, the engineering of FAPs provided the understanding of structure-function relationships of photoenzymes. The study of the novel BVMOs was the first identification and characterization in species *Cupriavidus necator*.

Zusammenfassung der Dissertation

Der Begriff „Biokatalyse“ bezeichnet die Nutzung von Enzymen oder ganzen Zellen als Katalysatoren für den Zugang zu einem breiten Spektrum chemischer Umwandlungen. Historische große Mängel (wahrgenommene oder reale) der Biokatalyse werden als Enzyme (Katalysatoren) bezeichnet, die teuer, instabil und ineffizient sind. Die Biokatalyse ist jedoch ein sehr dynamisches Forschungs- und Entwicklungsgebiet. Viele der bisherigen Mängel der Biokatalyse sind aufgrund der Fortschritte in der Molekularbiologie und Biotechnologie bereits behoben. Der Kern des Biokatalyse-Engineering besteht darin, Enzyme mit verbesserter Leistung (z. B. hohe Aktivität, gewünschte Substratspezifität, hohe Enantioselektivität und hohe Stabilität) zu entdecken und zu entwickeln. Die Verbesserung von Biokatalysatoren kann zur erfolgreicherer Anwendungen in Forschung und Industrie beitragen. Diese Dissertation befasste sich mit der Verbesserung von Biokatalysatoren für zwei verschiedene Enzyme, nämlich die Fettsäure-Photodecarboxylase (FAP, EC 4.1.1.106) und die Baeyer-Villiger-Monooxygenase (EC 1.14.13.x).

In Kapitel I wurden durch die Rekonstruktion der möglichen Vorfahren-Decarboxylasen aus einem Satz von 12 vorhandenen Fettsäure-Photodecarboxylase (FAP)-Sequenzen drei Vorfahren-Photodecarboxylasen ermittelt. Eines dieser Enzyme (ANC1) zeigte eine höhere Stabilität und Löslichkeit, aber eine geringere katalytische Aktivität als das vorhandene FAP. ANC1 weist mehr hydrophile Aminosäuren auf der Proteinoberfläche auf, was die Löslichkeit des Enzyms verbessern könnte. Die höhere Stabilität und geringere katalytische Aktivität könnte darauf zurückgeführt werden, dass es für ANC1 starrere Regionen gibt.

Eine weitere Verbesserung der Stabilität von CvFAP durch rationales Design und ortsgerechte Mutagenese führte zu einer thermostabilen Variante von CvFAPS473G. Der Rest E550 von CvFAP spielt eine entscheidende Rolle bei der katalytischen Aktivität. Die Variante sANC1V366L verbesserte die spezifische Aktivität als sANC1WT. Durch die Einführung von stark eutektischen Lösungsmitteln (DES) als alternatives Lösungsmittelsystem konnte die katalytische Aktivität für FAPs aufrechterhalten und verbessert werden.

In Kapitel II wurden in einem Screening 18 mutmaßliche BVMOs, aus dem Genom des fermentativen Bakteriums *Cupriavidus necator* H16 identifiziert, das auf natürliche Weise effiziente NADH-Regenerations-Dehydrogenasen enthält. Eines der BVMOs ermöglicht die Durchführung der BV-Oxidations- und Sulfoxidationsreaktion. Durch die Überexpression von BVMO in *C. necator* H16 benötigt die ganzzellvermittelte BV-Oxidation kein externes NADH. Beim Fluoreszenzassay zeigte dieses neuartige BVMO eine relativ hohe Schmelztemperatur.

Das Ziel dieser Dissertation ist es, die gewünschten Eigenschaften von Biokatalysatoren durch die Entdeckung und Entwicklung von Proteinen zu verbessern. Da das Photoenzym in der Natur selten vorkam, lieferte die Entwicklung von FAPs im Rahmen dieser Arbeit

das Verständnis der Struktur-Funktions-Beziehungen von Photoenzymen. Die Untersuchung der neuartigen BVMOs war die erste Identifizierung und Charakterisierung der Art *Cupriavidus necator*.

Danksagung

Firstly, I would thank Prof. Dr.rer.nat. Robert Kourist to give me this precious chance be a number of his group and his student. Everything just happened yesterday, when I have an interview with Prof. Dr.rer.nat. Robert Kourist. At that moment, we were talking about projects for Cyanobacteria. Robert just said we could use water and light, (snap), then get what we need. At that moment, I felt that is the enthusiasm for science. With three years of work with Robert, I learned the scientific knowledge and the passionate spirit to chase the dream. I am trained not just how to be a scientist but also how to tackle the obstacles for life. I am fortunate that I could learn with Robert. Once, when we have a group meeting, Robert just told us, hey, it is ok you make mistakes, you are here for learning, correcting you is my job as supervisor. So many details flash back in my memory. Learn with Robert was my honor all the time.

Next, I would thanks Assoc. Prof. Harald Pichler for being on my thesis committee and gave excellent suggestions for the projects. Also, thanks for his kind support for the documents and processes of the doctoral school. I would also thank for Asst. Prof. Regina Kratzer for being my thesis committee and for all the lovely suggestions for the projects. Thanks to Dr. Elia Calderini for contributing so much to the projects, from awesome ideas to practical suggestions. I am grateful for the scholarship from the Chinese Scholarship Council, which provide me the support to live and study in Austria.

I would thank my fantastic colleagues. Ivana, Sandy, Christoph, Lenny, Hanna, Anna, Leen, Jelena, Melissa, Feyza, Elif, Andrea, for all the friendly support. I would also thank Maria, Clemens, Simone, Daniel, Leonie, Amneris, and Svetlana for supporting the laboratory and everyone else. Many thanks to Eveline, Dani and Eva for the important organizational work in Office. So many thanks for our current and former group numbers, being a number of our lovely international group, would be my pleasant experience and memory always.

Last, I would thank my mum for all the love she gives me, without her, I would not make it today. I grow up in China, as a girl from a single-parent family, life was not easy for my mum and me. My mum worked as a worker who only makes a tiny salary, barely enough to survive. The darkest time for us was I started college, my mum had to take care of my sick grandma. Meanwhile, she had to work as a cleaning lady for a part-time job to support me. I could not forget when she cleans the street on the hottest summer day. My mum never complains anything to me. She always laughs and appreciates everything, even it is trouble. I always tell my friends to call me Sunny, because my mum shows me by action how to be a sunny person in life and bring sunshine to others. Yes, I have a superwoman in my life, Xing Jing (邢静) is her name.

Thank you.

The work presented in this thesis was conducted under the supervision of Univ.-Prof. Dr.rer.nat. Robert Kourist between December 2017 and May 2021 at the Institute of Molecular Biotechnology at Graz University of Technology.

Contents

Abstract of dissertation	I
Danksagung	IV
Chapter I	
Engineering functional thermostable Photodecarboxylases using ancestral sequence reconstruction.....	1
1. Abstract.....	2
2. Introduction.....	4
2.1. Biofuels	4
2.2. Biocatalysts for biofuels Alka(e)nes synthesis.....	7
2.3. Photodecarboxylase-CvFAP	11
2.4. Biocatalyst improvement by protein engineering.....	17
2.5. Reaction engineering- Deep Eutectic Solvents (DESs).....	18
3. Results and Discussion	19
3.1. Ancestral sequence reconstruction (ASR) of photodecarboxylase.....	19
3.2. Homology modeling and structural considerations.....	22
3.3. Heterologous expression and purification of FAP in <i>E.coli</i>	22
3.4. Study of FAPs expression level and activity by IPTG concentration or removing predicted chloroplast sequence	23
3.5. Photoinactivation of fatty acid photodecarboxylases.....	26
3.6. ANC1 and sANC1 have higher T _m and thermostability.....	30
3.7. Preparative-scale production of pentadecane	36
3.8. Influence of the putative chloroplast target sequence on the selectivity of ANC1.	37
3.9. Applicability of FAPs in deep eutectic solvents	37
3.10. Protein engineering of CvFAP stability and sANC1 catalytic activity.....	44
3.11. Rapid Nile red quantification method for pentadecane production.....	50
4. Conclusion and outlook	55
5. Experimental section	57
Chapter II	
Discovery and characterization of a Baeyer-Villiger Monooxygenase (BVMO) from <i>Cupriavidus necator</i> H16.	79
1. Abstract.....	80

2. Introduction	81
2.1. Oxygenases	81
2.2. Flavin-dependent monooxygenases.....	82
2.3. Baeyer-Villiger monooxygenases (BVMOs, E.C. 1.14.13.xx).....	83
2.4. <i>Cupriavidus necator</i> H16.....	87
3. Results and discussion	89
3.1. Discovery of Baeyer-Villiger monooxygenases from <i>C.necator</i> H16.	89
3.2. Determination of substrate scope for CPMO0.....	93
3.3. <i>In vivo</i> conversion of Cyclopentanone to δ -valerolactone by CPMO0.....	95
3.4. Heterologous expression, purification, and activity study of CPMO0 in <i>E.coli</i> . 95	95
3.5. Overexpression and purification of CPMO0 in <i>C.necator</i> H16.	101
3.6. <i>In vitro</i> activity assessment of purified enzyme of CPMO0.....	102
3.7. Melting Temperature determination of CPMO0 by ThermoFAD.....	103
4. Conclusion and outlook	106
5. Experimental section	107
References	141
Appendix A Sequence alignment of BVMOs	165
Appendix B The nucleotide encoding regions in <i>C.necator</i> H16.....	156
Appendix C Abbreviations	193
Appendix D List of Figures.....	200
Appendix E List of Tables	203

CHAPTER I

Engineering functional thermostable Photodecarboxylases using ancestral sequence reconstruction.

Chapter published in ChemBioChem
<https://doi.org/10.1002/cbic.202000851>

1. Abstract

“Light brings us the news of the Universe.”— William Henry Bragg

Very recently, a new fatty acid photodecarboxylase (FAP, EC 4.1.1.106) was reported. This photoenzyme belongs to an algae-specific clade of the glucose-methanol choline (GMC) oxidoreductase family, which comprises a wide range of oxidases and dehydrogenases. FAP contains a light-absorbing flavin adenine dinucleotide (FAD) cofactor and enables catalyze the blue light-driven ($\lambda=450\text{nm}$) decarboxylation of acid materials to C1-shortened alka(e)nes. CvFAP from *Chlorella variabilis* NC64A is the most extensively studied photodecarboxylase. It appears particularly attractive due to its capacity to synthesis alka(e)nes from fatty acids and the kinetic resolution of α -functionalized acids on the activation by light. However, the relatively low soluble expression level, production yield, and stability make it a challenging enzyme for research and application.

In this part of the thesis, we presented few basic investigations of FAPs, such as their catalytic activity and folding efficiency influenced by the light exposure, the plasmid backbone used, and the concentration of the inducer. Besides these basic investigations of FAPs, we achieved two major objectives about FAP:

1. Catalyst improvement.

We reconstructed possible ancestral decarboxylases from a set of 12 extant sequences that were classified under the fatty acid decarboxylases clade within the glucose-methanol choline (GMC) oxidoreductase family. One of the resurrected enzymes (ANC1) showed activity in the decarboxylation of fatty acids. ANC1 has a 15°C higher melting temperature (T_m) than the extant CvFAP. Its production yielded 12-fold more protein than wild-type decarboxylase, which offers practical advantages for the biochemical investigation of this photoenzyme. With rational design, CvFAP_{S473G} improved the stability of CvFAP, which has a 10°C T_m than the extant CvFAP. The residue E550 was identified as the critical position for CvFAP activity. The variant sANC1_{V366L} presented 3-fold specific activity than sANC1_{WT}.

2. Reaction engineering.

The exploration of FAP-mediated reaction using deep eutectic solvents (DESs) not only demonstrated in the DES reaction system, the catalytic activity for FAPs could be maintained and improved, also pave the way for practical applications of FAPs as a catalyst to produce a high yield of valuable chemicals with the simple process. Another result in this study was generating a fluorescence-based assay for detecting and quantifying pentadecane production. The fluorescent dye Nile red with palmitic acid and pentadecane presented different fluorescence characters. The fluorescent intensity could be used to evaluate the concentration of pentadecane. The assay could be used in whole cell and cell free extract systems. We have not optimized our method for efficiency, but

the procedure is significantly faster than analysis by GC. Using this assay, we would screen mutations with better capacity in a large number of mutations for a short period.

2. Introduction

2.1. Biofuels

Industrialization started in the north of England in the 19th century, the fossil fuels as energy sources have enabled society to grow and develop continuously for almost 200 years. During this period, coal was mainly the source at the beginning. Later various derived fuels from crude oil such as diesel, petrol, or kerosene were included. With the growth of economies and population, more fossil fuel combustion is consequently needed for more energy. The combustion of fossil fuels produces mainly two issues: 1, the product from the combustion process, carbon dioxide (CO₂) enhanced green-house effect and associated planet warming.^[1] 2, the declining production volume and the increasing prices of crude oil limited global economic growth.^[2] For tackling these two closely linked problems, moving away from dependence on conventional fossil fuel energy sources is necessary. There are multi-faceted approaches to mitigate the demand for fossil fuels, such as nuclear, solar, hydrogen, wind, geothermal, and biofuels. Among these options, biofuels are most promising in the short term as their market maturity.^[3]

Biofuels have emerged as a readily available energy alternative for conventional fossil fuels.^[4] They are referred to energy enriched chemicals generated via biological processes or derived from the biomass of living organisms (microalgae, plants, and bacteria). For example, ethanol, methanol, biodiesel, Fischer-Tropsch diesel, hydrogen, and methane are the fuels produced from biomass.^[5]

Renewable and carbon-neutral biofuels are important because they could replace petroleum fuels necessary for environmental and economic sustainability. Between 1980 and 2015,^[6] worldwide production of biofuels increased by a magnitude order, from 4.4 to 50.1 billion liters, and with promising further dramatic increase.^[7]

Generally, biofuels are classified as primary and secondary biofuels. The primary biofuels are the natural and unprocessed form, such as fuelwood, wood chips, and pellets, etc. primary fuels are directly combusted for heating, cooking, or producing electricity. The secondary fuels are the products from the biomass process, e.g., ethanol, biodiesel, DME, etc. Depending on the kind of raw material and technology used for production, the secondary biofuels are divided into first, second, and third-generation biofuels. (Figure 1) Secondary biofuels can be used for multiple ranges of applications, like transport and high-temperature industrial processes.

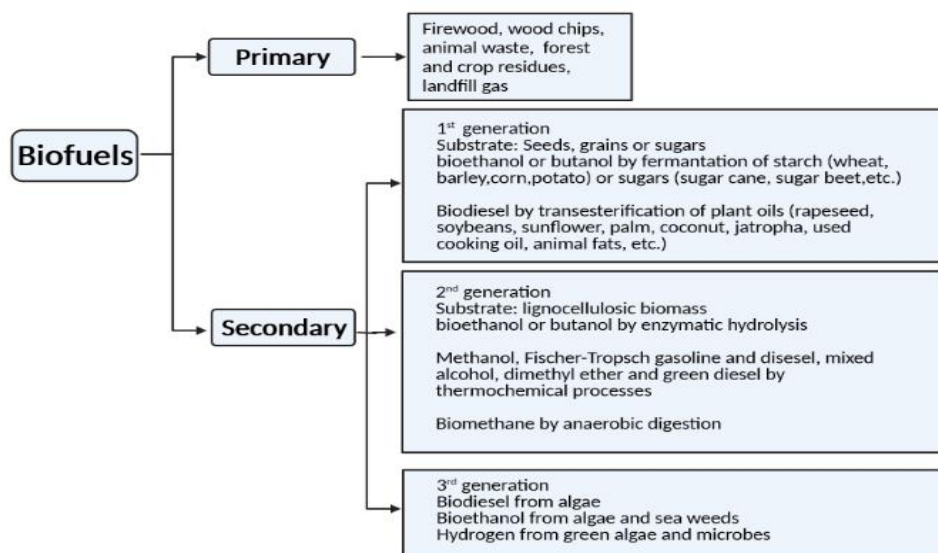


Figure 1. Classification of biofuels.

The first-generation biofuels are produced with a relatively simple process from sugars,^{[8]-[10]} grains, or seeds.^{[11],[12]} Ethanol is the most well-known first-generation biofuels made by fermentation of sugar extracted from crop plants or starchy crops.^{[9],[13]} Another popularized first-generation biofuels is biodiesel produced by transesterification processes from vegetable oils of oleaginous plants.^[13] The critical challenge for the viability of first-generation biofuels is the conflict with the food supply.^[14] Thus, searching for non-edible biomass for producing biofuels will be necessary.

The second-generation biofuels from non-food feedstock lignocellulose address some of the problems from first-generation fuels, such as lower the cost of feed material, greater production, and higher land using efficiency. However, two serious obstacles for using low-price lignocellulosic feedstocks as substrates to produce fuels: 1, the polymeric cellulose, hemicellulose, and lignin must be broken down biochemically or thermally into simpler compounds (e.g., sugars, organic acids); 2, the monomers or oligomers after pretreated processes from the first step need to be deoxygenated into biofuels.^[13] From an economically competitive opinion with fossil fuels, these two challenges heavily influence the ultimate cost of second-generation biofuels.

The most standard definition for third-generation biofuels is the fuels be produced from algal (microalgae and macroalgae) biomass, which being devoid of major drawbacks associated with first/second-generation biofuels. Algae are present in all existing ecosystems as one of the oldest life forms.^[4] They are primitive plants living in a wide range of environmental conditions.^[15] Phototrophic algae could absorb sunlight and carbon dioxide from air and nutrients from aquatic environments.

Microalgae can produce lipids, proteins, and carbohydrates, which can be processed into biofuels and valuable products (Figure 2).^[16] The utilization of photosynthetic organisms as a preferred source for biofuels is cheap and reliable.^[17] For example, *Synechococcus*

elongatus PCC 7942 presented high production for isobutyraldehyde and butanol.^[18] Some species of green algae, *Botryococcus braunii* and *Chlorella protothecoides* could produce shorter hydrocarbons from terpenoid hydrocarbons and glyceryl lipid in their biomass.^[19]

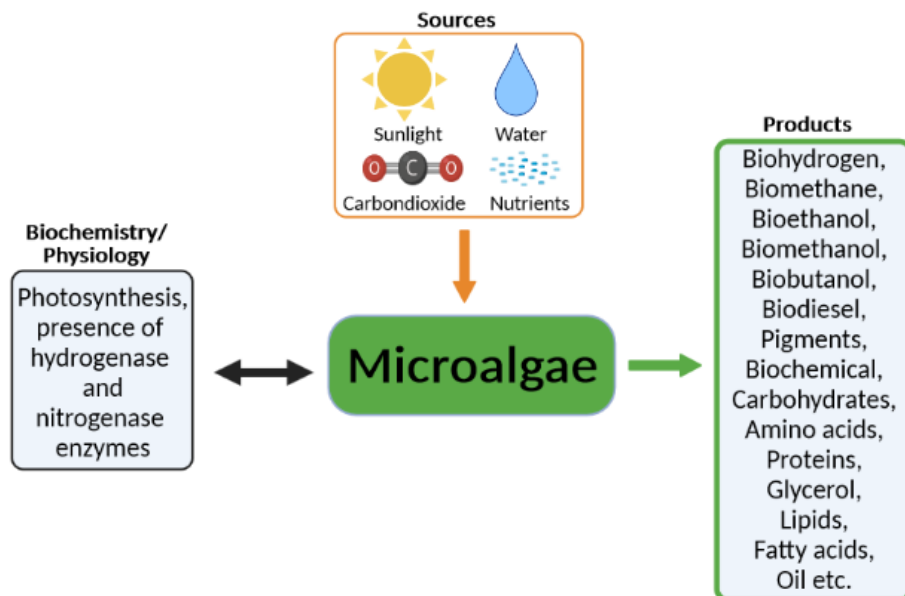


Figure 2. The working mechanism of microalgae in the energy sector.

To produce the ideal advanced fuel, few factors need to be considered, energy content, combustion quality, volatility, lubricity, viscosity stability, toxicity, water miscibility, and cost et al. Some biofuels exhibited similar properties to petroleum-based fuels (Table 1).^{[20],[21]} However, biofuels still have several drawbacks, for example, bioethanol from its biochemical properties, the lower energy density than petroleum-derived fuels, the hygroscopicity and volatility make the storage difficult.^[22]

Table 1. Conventional transportation fuel properties and potential biofuels

Fuel type	Major components	Properties	Biosynthetic alternatives
Gasoline	C ₄ -C ₁₂ hydrocarbons Linear, branched, cyclic aromatics Anti-knock additives	Octane number (87-91) Energy content Transportability	Butanol, isobutanol, short-chain alcohols, short branched-chain alkanes
Diesel	C ₉ -C ₂₃ hydrocarbons Linear, branched, cyclic aromatics Anti-freeze additives	Cetane number (40-60) Low freezing temperature Low vapor pressure	Fatty alcohol, alkanes, linear or cyclic isoprenoids
Jet fuel	C ₈ -C ₁₆ hydrocarbons Linear, branched, cyclic aromatics Anti-freeze additives	Net heat of combustion Density Very low freezing temperature	Branched alkanes, linear or cyclic isoprenoids

2.2. Biocatalysts for biofuels Alka(e)nes synthesis

As one of the hydrocarbon biofuels, alka(e)nes constitutes most petroleum-derived fuels (gasoline, diesel, and jet fuels).^[23] On the other hand, the key features of alka(e)nes (chain length; degree of saturation; straight or branched) contributed to their wide utilization. Alka(e)nes are the ideal organic compounds for combustion and basic components for producing various materials.^[24] The observations of alka(e)nes production from plants, insects, microbes, and microalgae started a long time ago.^{[25]–[28]} In higher plants, alkanes are mainly involved in synthesizing the epicuticular wax layer.^[29] The role of hydrocarbons in microorganisms is still not completely clear.^[30] Many efforts have been made to produce alka(e)nes in engineered microbial strains (Figure 3). Integration of metabolic engineering and synthetic/system biology resulted in remarkable engineering microorganisms as cell factories to produce alkanes/alkenes for industrial production.

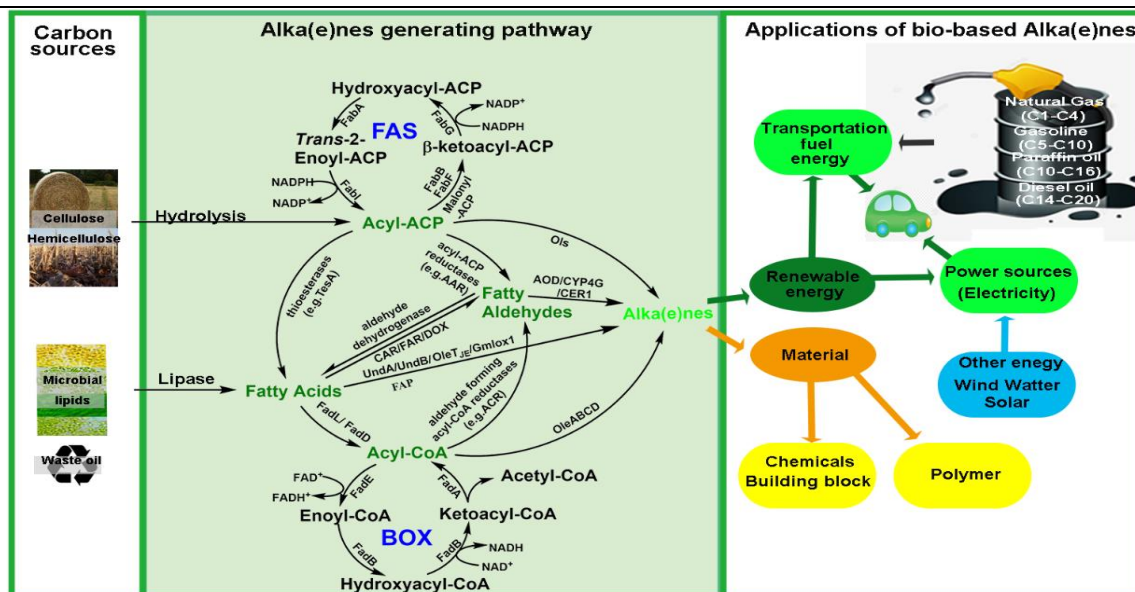


Figure 3. Contribution of bio-based hydrocarbon from low-value substrates via metabolic engineering.

Figure adapted from Wang *et al.*^[31] FAS, fatty acid synthesis, Box, β -oxidation. There are four major precursors (fatty aldehydes, fatty acid, acyl-ACP, and acyl-CoA) for alka(e)nes biosynthesis.

As shown in Figure 4, for the final step to synthesis alka(e)nes, the studies mainly focus on six types of alka(e)nes generating pathways: a, Thiolase-based system (OleABCD); b, Olefin synthase (Ols); c, Olefin-forming fatty acid decarboxylase (OleT_{JE}); d, Nonheme mononuclear iron oxidase UndA and UndB; e, Aldehyde deformylating oxygenase (ADO); f, Fatty acid photodecarboxylase (FAP).

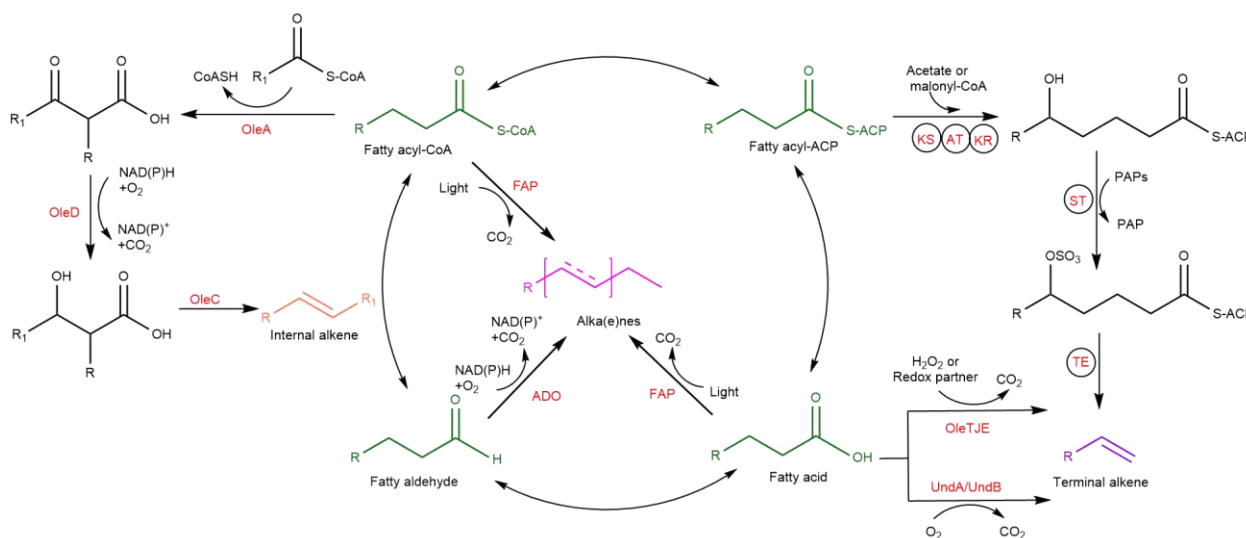


Figure 4. An overall diagram of enzymatic reaction for converting precursor substrates to alka(e)nes.

a) Thiolase-based system (OleABCD)

For enzymes involved in the long-chain internal alkene (C₂₂-C₃₃) biosynthesis, OleABCD enzymes were the combination of a superfamily of enzymes consisting of thiolase (OleA), α/β -hydrolase (OleB), AMP-dependent ligase/synthase (OleC), and dehydrogenase/reductase (OleD).^[32] The OleA initiates the proposed pathway by generating a β -keto acid, then OleD produces β -hydroxy acid by NADPH-dependent reduction.^[33] OleC then converts the β -hydroxy acid to alkene by consuming ATP.^[34]

b) Olefin synthase (OLS)

Generally, OLS is a modular multi-domain synthase, consisting of three modules (activation module, extension module, terminal alkene formation module).^[35] The activation module locates at the N-terminal of OLS is the first modular, comprised of fatty acyl-ACP ligase (FAAL). The extension module comprised three domains, ketosynthase (KS), acyltransferases (AT), ketoreductase (KR).^[36] The catalytic processes initiate from FAAL with ATP to yield fatty acyl-ACP intermediate, which the acyl chain be elongated by extension module. The KS and AT add two carbons from malonyl-CoA to fatty acyl-ACP intermediate. After elongation, the β -carbonyl group is reduced to the β -hydroxyl group by KR. The next step is the decarboxylation and dehydration catalyzed by the TE domain to release the terminal alkene product. This distinct biosynthetic pathway was identified in cyanobacteria only related to long-chain terminal olefin (predominantly C₁₉).^[37]

c) Olefin-forming fatty acid decarboxylase (OleT_{JE})

OleT_{JE} is an attractive enzyme attributed to its ability to produce terminal alkenes from fatty acid in a single step. OleT_{JE} is a cytochrome P450 enzyme that catalyzes medium- to long-chain fatty acid (C₁₂-C₂₀) into 1-alkenes (C_{n-1}).^[38] The decarboxylation process of OleT_{JE} needs H₂O₂ as electron donor and oxidant. Further study showed OleT_{JE} also enables H₂O₂-independent redox partner systems: RhFRed, Fdr/Fdx, and CamAB^[39] for 1-alkene production.

d) Nonheme mononuclear iron oxidase UndA and UndB

The decarboxylases UndA and UndB were discovered from *Pseudomonas* species.^{[40],[41]} The identification of this non-heme oxidases/decarboxylases group initiated with UndA, which specifically converts medium-chain fatty acids (C₁₀-C₁₄) to their corresponding C_{n-1} 1-alkenes.^[40] Further biochemical and structural analyses demonstrated UndA requires Fe²⁺ for activity. This enzyme showed low activity both *in vitro* and *in vivo*, the reasons still unclear. UndB enables converted medium-chain fatty acids (C₆-C₁₆) to their corresponding C_{n-1} terminal olefins. Different from UndA, UndB is a membrane-bound desaturase-like enzyme.

e) Aldehyde deformylating oxygenase (ADO)

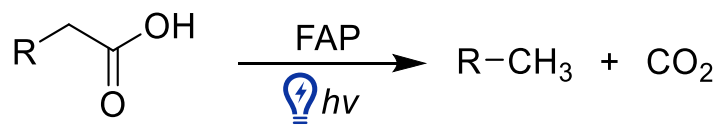
Aldehyde decarbonylases (ADs) have been discovered from various organisms, such as plants, insects, and cyanobacteria.^{[42],[43]} Due to its reliability of ADO in a different host, most metabolic engineering or cascade reactions for the produce of alka(e)nes were

based on fatty aldehydes.^[44] ADO is a non-heme di-iron enzyme that catalyzes the decarbonylation of fatty aldehydes requiring oxygen as an oxidant and needs a reducing system as an electron donor (NADPH; ferredoxin, and ferredoxin reductase).^[45]

f) Fatty acid photodecarboxylase (FAP).

Very recently, a new fatty acid photodecarboxylase (FAP, EC 4.1.1.106) was reported.^[46] This photoenzyme belongs to an algae-specific clade of the glucose-methanol choline (GMC) oxidoreductase family, which comprises a wide range of oxidases and dehydrogenases. FAP contains a light-absorbing flavin adenine dinucleotide (FAD) cofactor and catalyzes the blue light-driven ($\lambda=450\text{nm}$) decarboxylation of acid materials to C1-shortened alkanes (Scheme 1). The identification of FAP complemented the routes in some microalgae, including *Chlorella variabilis* NC64A and *Chlamydomonas reinhardtii* 137C, to convert fatty acid to alka(e)nes under the light.^[47]

After FAP was found, it caused some attention immediately attributing to its outstanding properties. Firstly, it exhibits turnover numbers in the range of thousands, and it was significantly higher than other alka(e)nes synthases, such as K_{cat} of ADO is only 0.36 min^{-1} . This suggested FAP would be feasible for biosynthesis alka(e)nes on a large scale. Secondly, comparing with other synthases, FAP only needs photons from blue light to activate its FAD cofactor for catalysis, and the only byproduct is CO_2 , which suggested FAP is ideal for eco-friendly and sustainable applications. Third, FAP exhibits the ability to transform $\text{C}_1\text{-C}_{20}$ fatty acids.



Scheme 1. General reaction scheme for FAP

Photoenzymes have seldom been found in nature, and there are two families of proteins bearing the light-drive activity that could be found in nature, the photolyases involved in the repair of DNA damage in many organisms^{[48]-[51]} and the light-driven protochlorophyllide reductases involved in chlorophyll synthesis.^[52] The chlorophyll f synthase was considered as the other type of photoenzyme, but the light-force nature of the reaction has not been established.^[53] Therefore, photodecarboxylases FAP was considered as the third novel type of photoenzyme.

2.3. Photodecarboxylase-CvFAP

Among FAPs, CvFAP is the first identified and characterized photodecarboxylase from the green microalgae *Chlorella variabilis* NC64A.^[46] Several recently published studies about CvFAP demonstrated its photocatalytic mechanism, catalytic activity, selectivity, capability to generating photoenzymatic cascades with additional enzymes, and applicability in the various microbial platform. (Figure 5).

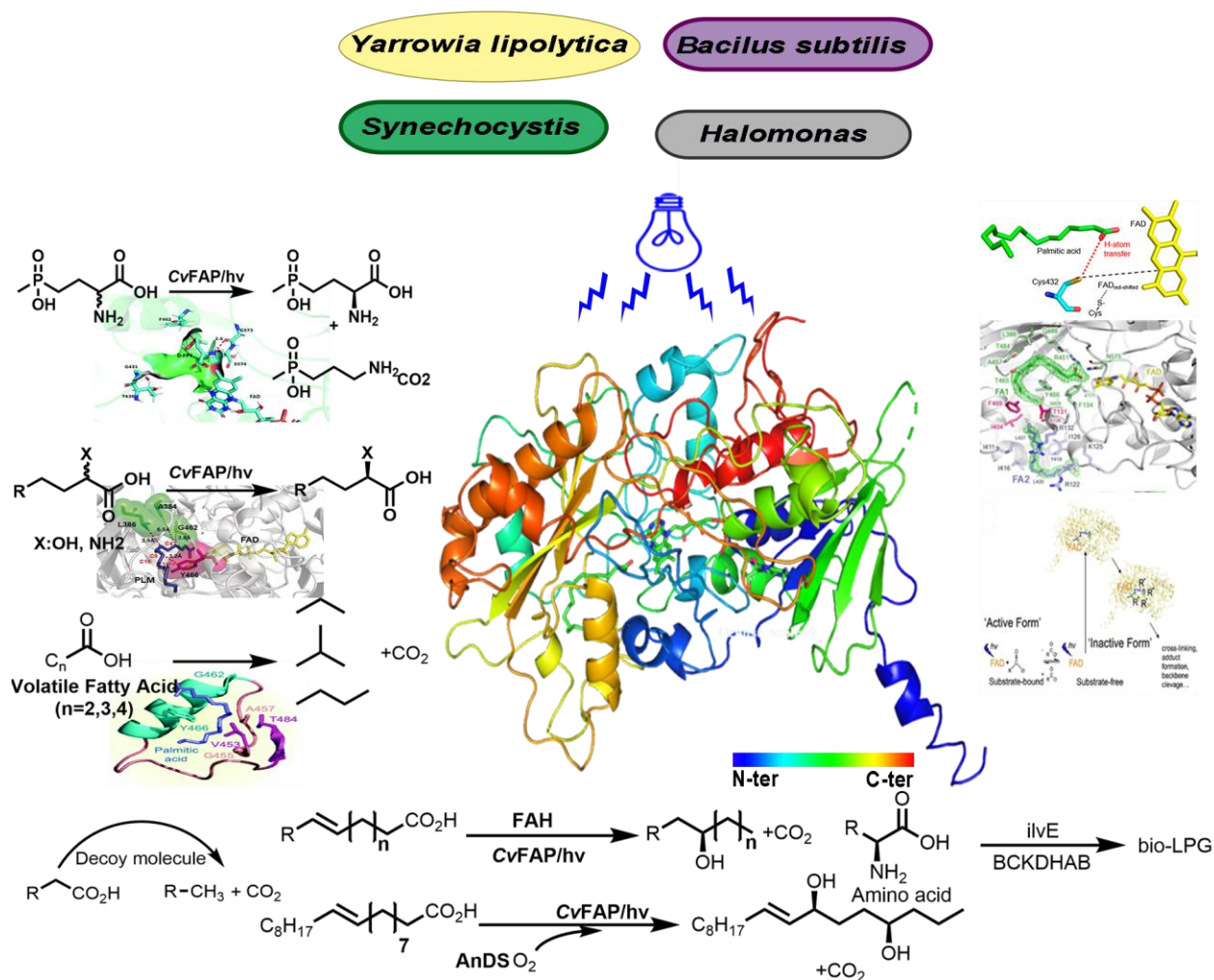


Figure 5. Recently published study about CvFAP.

Firstly, no matter from fundamental research requirements or application point of view, understanding the photoenzymatic mechanisms of FAP in detail is of utmost importance. The suggested photocycle (Figure 6) in FAP initiates with photoexcited flavin (FAD^*) abstracts one electron from fatty acid substrates to form the flavin semiquinone ($\text{FAD}^{\bullet-}$) and an alkyl radical on fatty acid (R^\bullet), this facilitates decarboxylation and release of carbon dioxide. Then the flavin semiquinone transfers back one electron to reduce the alkyl radical to form alkanes (RH). The reoxidation of FAD semiquinone results in a transiently red-shifted flavin state ($\text{FAD}_{\text{red-shifted}}$), which reverts to the initial state of flavin (FAD). For the two strictly conserved residues Arg_{451} ^[54] and Cys_{432} ^[55], two studies identified as proton donors. However, these two residues add proton/hydrogen atom and transfer electron remain to be clarified. One part of Scrutton's group work is also involved in photoenzymatic mechanisms study. Their study demonstrated the blue light exposure irreversibly inactivated the CvFAP enzyme since the formation of protein-based organic radicals.^[56]

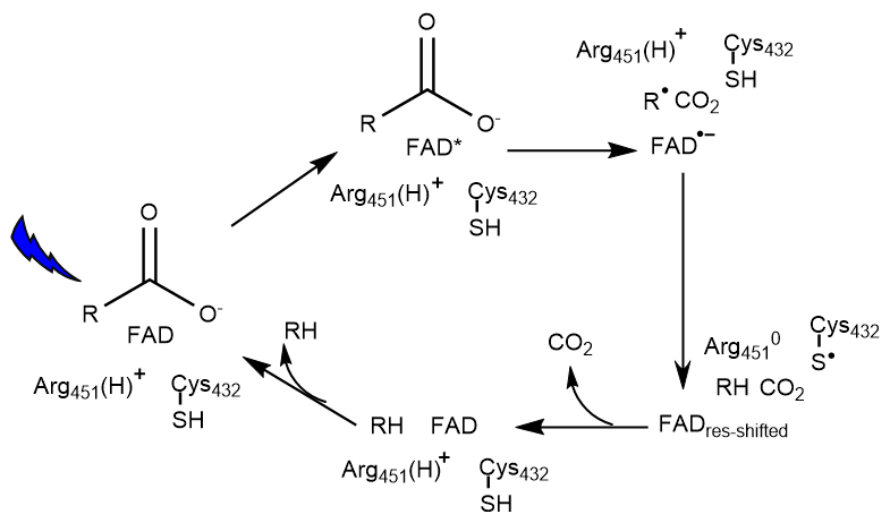


Figure 6. Suggested CvFAP photocycle.

The bienzymatic cascade comprising lipase from *Candida rugosa* and CvFAP was the first example of photoenzymatic cascade (Figure 7) which involved photodecarboxylase. This two-step cascade reaction was CvFAP-catalyzed photodecarboxylation of the free oleic acid from triolein, which was hydrolyzed from lipase.^[57]

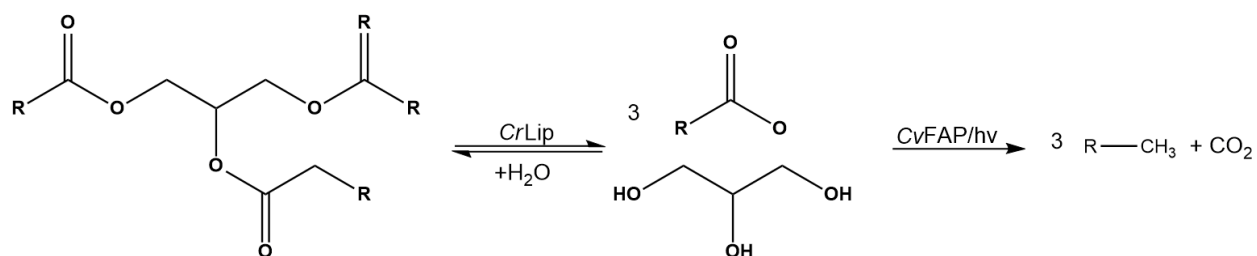


Figure 7. Photoenzymatic cascade to transform triolein into (Z)-heptadec-8-ene.

Scrutton's group initially attempted to convert butyrate by eight cyanobacterial FAP homologs. The results demonstrated that all FAPs presented catalytic activity, and the highest propane levels were detected with CvFAP.^{[58],[59]} Latter improved the production of bio-LPG (bio-propane and bio-butane) via rational design engineering of CvFAP. Since the feedstock volatile fatty acid (VFA) is inconvenient to manipulate, the precursors of VFA, amino acids (valine, leucine, and isoleucine) were selected to synthesis bio-LPG.^[60] Two photoenzymatic pathways comprising multiple additional enzymes and the CvFAP variant (Figure 8) were constructed. Transformation of VFA from amino acid is the first step. VFA could be generating from two routes, which are KdcA-dependent and CoA-dependent pathways. The leucine 2-oxoglutarate transaminase *ilvE* from *E. coli* generates deamination of amino acids yielding α -keto acids. Then in CoA-dependent route contain the α -keto acid dehydrogenase complex (BCKDHAB)^[60] and acyl-CoA thioester hydrolase from *Haemophilus influenza* (*YciA*)^[61], in the KdcA-dependent route, contain the decarboxylase (*KdcA*) from *Lactococcus lactis*^[62] and AIDH enzymes.^[60] These two photocatalytic pathways for the production of bio-LPG were transited into *E.coli*, *Halomonas*, and *Synechocystis*.

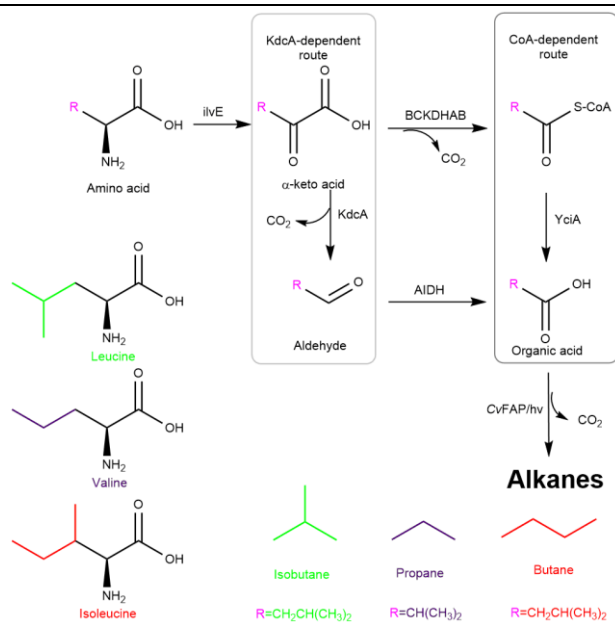


Figure 8. The overall scheme of the biocatalytic production of gaseous hydrocarbons from amino acids.

The photoenzymatic cascade reaction combined fatty acid hydratases (FAHs) (Figure 9, cascade 1) from *Lactobacillus reuteri* or *Stenotrophomonas maltophilia* with CvFAP converted a broad range of unsaturated fatty acids into enantiomerically pure secondary fatty acid alcohols. Also, the cascade combined 5,8-diol synthase (Figure 9 cascade 2) from *Aspergillus nidulans* (AnDS) with CvFAP yielded diols.^[63] In the first step, the hydratases stereoselectively (di)hydrated the C=C double bond. The following step is CvFAP catalyzing the photodecarboxylation of the intermediate hydroxyl acid. These two cascades successfully synthesize the long-chain secondary alcohols, which are currently not accessible from natural fatty acid through a biological approach.

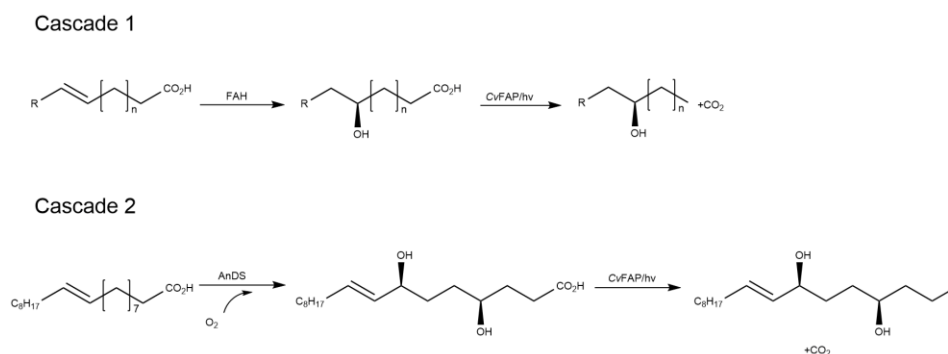


Figure 9. Photoenzymatic cascades to transform unsaturated fatty acid into secondary alcohols.

Multifunctional chiral molecules, such as α -amino acids and α -hydroxy carboxylic acids, are essential building blocks in pharmaceutical chemistry and chemical biology. Wu's group described the first CvFAP-catalyzed enantioselective example.^[64] They introduced

larger amino acids through site-directed mutagenesis at substrates and the binding tunnel of CvFAP. The CvFAP variants catalyzed the kinetic resolution of α -functionalized amino acid and carboxylic acids, which provided the unreacted *R*-configured substrates with high yields and excellent stereoselectivity.

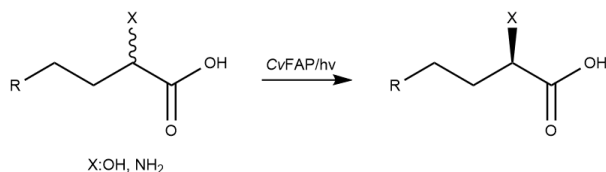


Figure 10. Asymmetric synthesis of α -hydroxy/amino acids by CvFAP.

Though semi-rational engineering method, CvFAP enables catalyzed deracemization of phosphinothricin (PPT) and provided unreacted L-PPT with high yields and excellent stereoselectivity.^[65] Since CvFAP is a light-driven enzyme, this novel synthesis approach required neither NADPH nor PPT precursors for transferring electrons, which was necessary for other approaches.

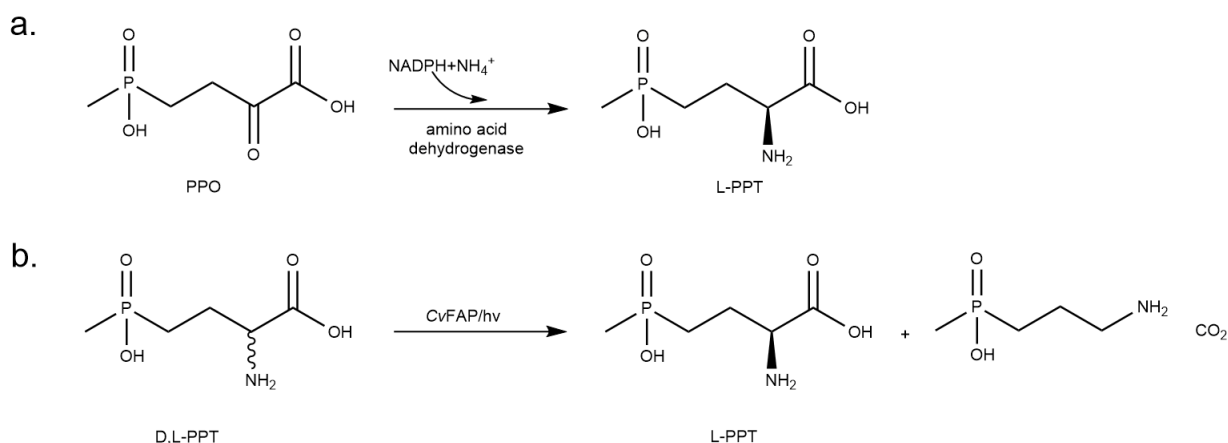


Figure 11. Enzymatic asymmetric synthesis of L-phosphinothricin (L-PPT).

- a. Bioasymmetric synthesis of L-PPT from 4-(hydroxymethylphosphinyl)-2-oxobutyric acid (PPO) via engineered dehydrogenase. b. CvFAP based photocatalytic deracemization of phosphinothricin (PPT).

Yarrowia lipolytica is one of the most widely recognized models of oleaginous organisms, sharing the same alka(e)nes synthesis pathway as cyanobacteria.^[66] As work from Stephanopoulos et al. (Figure 12), they report the acyl-CoAs being the preferred substrate for CvFAP rather than a free fatty acid. From this newly discovered pathway, the higher-titer alka(e)nes could reach 1.47 g/L from glucose.^[67]

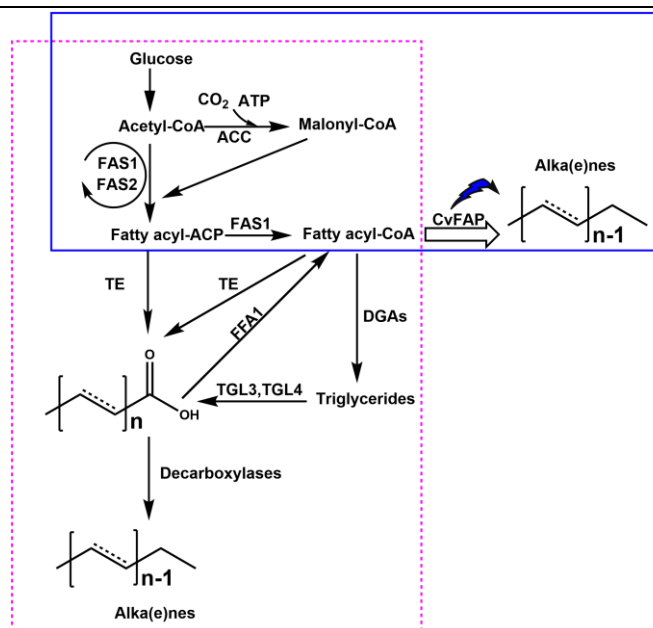


Figure 12. Metabolic pathways for alka(e)ne synthesis in *Yarrowia lipolytica*.

The pink dashed square represents the heterologous pathways employing the fatty acid decarboxylase according to previous literature^[66] that free fatty acids are the precursors of alka(e)nes. The blue square represents the revised pathway for alka(e)nes production using fatty acyl-CoA as precursors for CvFAP. TE, acyl-CoA/acyl-ACP thioesterases; TGL3 and TGL4, intracellular lipases from *Y. lipolytica*; FAA1, acyl-CoA synthase, ACC, acetyl-CoA carboxylase, FAS1 and FAS2, fatty acid synthase 1 and 2, DGAs, diacylglycerol transferases.

Displaying the enzymes on the surface of spores is the rapid and simple biotechnological production of immobilized enzymes. Therefore, Kabisch's group developed the *Bacillus subtilis* spore as the platform to produce hydrocarbons. In the bienzymatic one-pot cascade, the reaction was completed by displaying lipase and CvFAP on the *Bacillus subtilis* spore surface.^[68]

CvFAP exhibits a clear preference for long-chain fatty acids (C₁₆-C₁₇). Hollmann's group provided the decoy molecule approach, which enables converting an extensive range of carboxylic acids by filling up the vacant substrate access channel of CvFAP.^[69] This approach enlarged the substrate scope of CvFAP not limited to straight-chain carboxylic acids, as it also could convert some branched and unsaturated carboxylic acids with effective rates.

Despite CvFAP's extensively recent discoveries, synthetic applications of CvFAP are limited by its poor thermal stability and sensitivity to solvents and light. All this results in a limited product formation, which cannot pay off the effort to produce the enzyme. On top of the low stability, production yields remain low, which is an obstacle for both the enzyme's biochemical characterization and industrial application. The study for improving the stability of CvFAP still was vacant.

2.4. Biocatalyst improvement by protein engineering

Ancestral sequence reconstruction (ASR) is the probabilistic reconstruction of ancient protein sequences based on extant genes. It has emerged as a proficient method to produce functional proteins that keep the aspects of the protein family.^{[70],[71]} Reconstructed enzymes may not only have a broader substrate range compared to their modern descendants^[72] but often show improved thermostability and solvent tolerance.^{[73],[74]}

Besides the ASR method, we also considered the general categories of enzyme engineering, include directed evolution,^[75] rational design,^[76] and semi-rational design.^[77] Directed evolution could improve the stability of enzymes by emulating natural evolution,^[78] even without requiring the comprehension of protein-function relationships. However, on the other side, efficient high-throughput screening or selection methods and enormous scope random mutant libraries are necessary. With sufficient information on protein structure, catalytic mechanism, and structure-function mechanism, rational design is an efficient strategy.^[79]

Compared with directed evolution and rational design, the semi-rational design combines the information of protein sequence, structure, and function with predictive algorithms to preselect promising target sites and restricted amino acid diversity. Therefore, the semi-rational design offers promising predictors for designing smaller, higher-quality libraries from a practical perspective. Based on bioinformatics analyses, the computer-aided approaches ranging from analyzing the primary sequence, simulating the tertiary structure, and *de novo* predicting the protein interaction.^[80] The growing availability of bioinformatics tools together with novel algorithms significantly improved the efficiency of prediction.

With the advent of molecular techniques, the increasing availability of sequences raised the popularity of bioinformatics. In the past years, the complexity of the algorithms has dramatically expanded the capacity to obtain more valuable information from sequence data.^[81] In addition, various algorithms have been created to predict the protein function based on the sequences and 3D structure of a protein^[82] to anticipate more stable mutations^[83] or change the enzyme's substrate specificity.^[84]

Rosetta software suite and HotSpot Wizard are generally efficient prediction tools for semi-/rational protein engineering. For the Rosetta software suite, the point mutant (pmut) scan application is the computational program that calculates the relative change in folding free energy $\Delta\Delta G$ for predicting protein stability. Theoretically, a more stable mutant has a negative $\Delta\Delta G$. We used the program RosettaDesign, which uses an energy function for evaluating the fitness of a particular sequence for a given fold, and a Metropolis Monte Carlo search algorithm for sampling sequence space. The program requires a backbone structure as input and generates sequences with the lowest energy for that fold. HotSpot Wizard is a web server used to automate hotspots in semi-rational protein design to improve protein stability, catalytic activity, substrate specificity, and enantioselectivity. HotSpot Wizard 3.0 could accept the protein sequence as input data.

As substrates for photodecarboxylase, the poor solubility of free fatty acids in water limit the load of the substrate, and furthermore, a co-solvent is always needed to increase the substrate solubility further. As we know, organic solvents can interact with the hydrophobic patches of proteins perturbing their folding/unfolding equilibrium in aqueous solutions.^[85] An alternative solvent system would be desirable.

2.5. Reaction engineering- Deep Eutectic Solvents (DESs).

Up to now, a new generation of available green solvents has emerged, so-called Deep Eutectic Solvents (DESs). This eutectic mixture could be formed by simply mixing a few cheap, renewable, and biodegradable components.^{[86],[87]} As a green medium, DESs could be obtained by the complexation of quaternary ammonium salts with hydrogen-bond donors^{[88],[89]} and meet different criteria such as availability, non-toxicity, biodegradability, recyclability, and low price.^[90] Several attempts were also made in biocatalysis where DES has been used as (co)-solvents catalyzed reactions.^{[91],[92]} Unlike in organic solvents, the enzymatic activity is dramatically reduced, DES and DES/water mixtures have been shown to preserve such activity partially. Moreover, concerning DESs and microorganisms, the stability of lyophilized *E. coli* was proven in DES.^[93]

In view that CvFAP catalyzes fatty acid substrates to generate alkenes, either fatty acid consumption or alkene production could be selected as the means to evaluate enzyme activity. The only readily detectable approach for fatty acid and alkanes was a highly specified and sophisticated GC approach, which needs complicated preparative steps, i.e., extraction, purification, and concentration. Consequently, the engineering study of catalytic activity or stability of FAPs can be extremely energy and time-consuming. A faster and more convenient method to accelerate the quantification process will be ideal.

Fluorescence spectroscopy is a simple, rapid, and sensitive detectable technique used as combinatorial and high throughput experimentation techniques. Nile red is a benzo phenoxazine dye with the properties of a fluorescent hydrophobic probe. In the presence of hydrophobic compounds (i.e., fatty acids, membrane lipids),^[94] the fluorescence character of the dye shifts depending on the hydrophobicity of the compounds and their surroundings.^[95] Such as Nile red with neutral lipids, the maximum emission wavelength is shorter, and the fluorescence intensity also higher than with polar lipids.^[95] Lipids composed of unsaturated fatty acids presented stronger fluorescence intensity than saturated fatty acids.^[96] As the results from *Spencer et al.*, they presented a high throughput screening for intracellular lipid droplets based on the hydrophobic fluorescence dye Nile Red.^[97] Moreover, the fluorescence intensity of Nile red scales with its hydrophobic environment^[98] was used to evaluate the lipid content of bacteria,^[99] microalgae,^[100] and yeast.^[94]

In our case, FAPs catalysis decarboxylation of fatty acids to produce alkanes. We reasoned that the Nile red with fatty acid and alkane would present different fluorescence characters. The fluorescent intensity would be used to evaluate the concentration of fatty acids and alkenes.

3. Results and Discussion

3.1. Ancestral sequence reconstruction (ASR) of photodecarboxylase

The typical first step of ancestral sequence reconstruction is collecting a set of recent homologous sequences. The most straightforward way to gather a large number of sequences is a similarity search. The major downfall is that public databases contain more sequences than needed for a successful ASR. In our case, a second risk was present as CvFAP is a member of the glucose-methanol-choline (GMC) superfamily. Thus, a BLAST search would produce a highly unbalanced dataset with the risk of losing the light-dependent activity and producing an ancestral oxidoreductase. Given the scarce frequency of photoenzymes, we decided to include only 12 sequences classified in the fatty acid photodecarboxylases clade by Sorigue and coworkers where only two enzymes had confirmed photodecarboxylation activity (Figure 13). All sequences belong to the algal genus, and we hypothesized that the ASR would generate a less organism-specific enzyme more compatible with the expression machinery in *E. coli*.

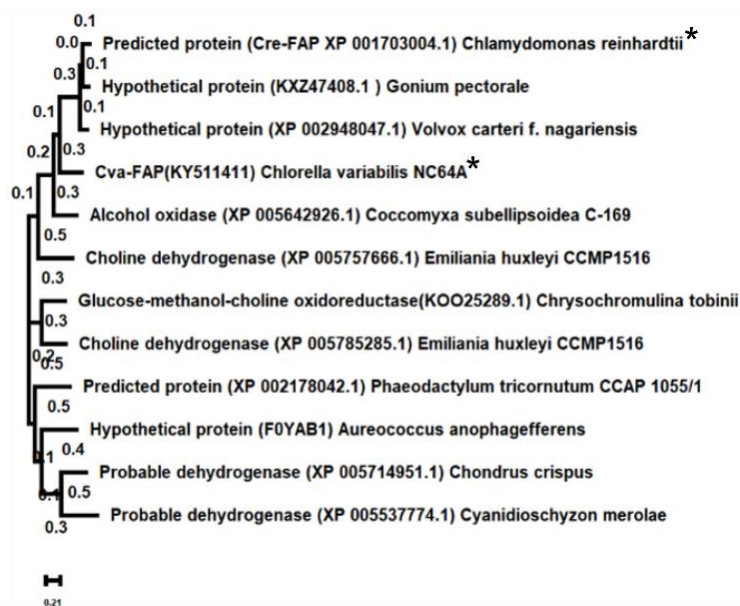


Figure 13. Phylogenetic relationship of the 12 sequences of extant FAP enzymes used to reconstruct the ancestral sequences.

* indicates enzymes with confirmed photodecarboxylation activity.

Aware of the limitation of the dataset, three different reconstructions named ANC1, ANC2, and ANC3 were designed by using different sequence alignment tools and evolutionary models to increase the probability of obtaining an active variant. The sequences of the node 0 elements from the three different reconstructions were ordered as codon-optimized synthetic genes. The sequence identity with the wild type was determined to be 76% for ANC1 and ANC2 and 54% for ANC3, respectively. Figure 14 visually illustrates how ANC1 and ANC2 are more similar to WT (more black bars), whereas ANC3 has a lower sequence identity (more white bars). Despite the differences, they all share the

conserved amino acid positions that characterize FAPs: C432, R451, and A576 (numbering corresponding to CvFAP) (Figure 15). The latter is the most striking difference between FAPs and other members of the GMC superfamily, as typically here, a histidine is present.

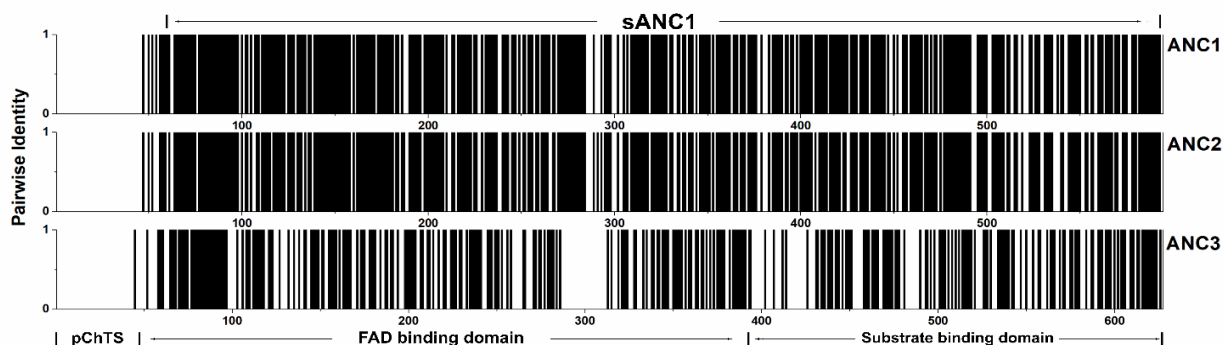


Figure 14. Pairwise identity between each ancestor and CvFAP.

The residue number is given in the X-axis, whereas the y-axis denotes pairwise identity at each site. Black bars denote the pairwise identity of 100% at each site. The portion of ANC1 corresponding to sANC1 is highlighted at the top of the figure. (pChTS: predicted Chloroplast Target sequence).

3.2. Homology modeling and structural considerations

The wild-type crystal structure (PDB-ID: 5NCC) is available, the homology models for all ancestors were generated by using the Rosetta server.^[101] Aligning the homology models to the experimental crystal structure showed few structural variations, primarily characterized by shorter loop regions, whereas ANC3 seems to have more differences than the CvFAP structure (Figure 16).

As the homology models were built based on the wild-type structure, no dramatic structural changes were visible; the most striking difference, especially in the active ANC1, is the length of few flexible loops. The flexible loops in the ancestor seem to be shorter, and this might have decreased the enzyme's flexibility increasing its stability.

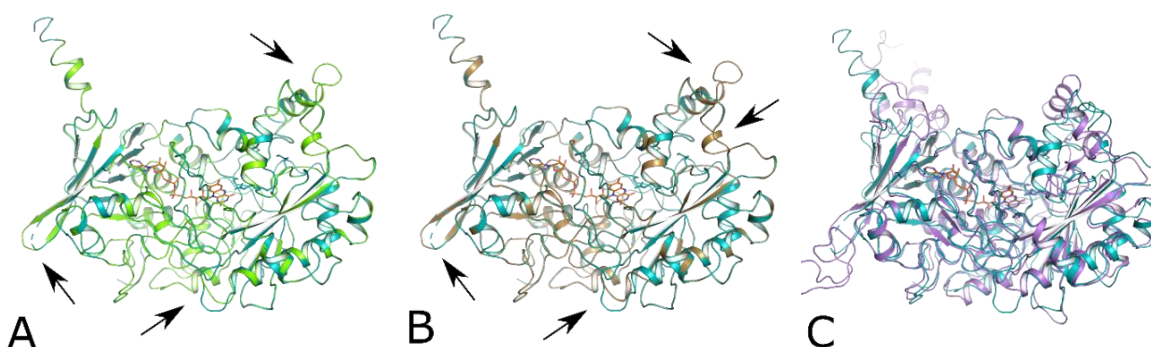


Figure 16. Homology models of ancestors superimposed on CvFAP structure (PDB-ID: 5NCC).

Black arrows highlight the main differences visible between WT and the ancestors' homology models. ANC3 in C appears significantly different. CvFAP is teal in A, B and C; ANC1 is green in A; ANC2 is sand in B; ANC3 is purple in C. Images generated in PyMOL 2.3.5.

3.3. Heterologous expression and purification of FAP in *E.coli*

Protein solubility and purification yields are important prerequisites for successful biotechnological studies and industrial applications.^{[72],[74]} CvFAP was previously expressed in *E.coli*, the soluble protein and purification yield were relatively low. Meanwhile, purification often leads to loss of catalytic activity.^[56] The production of pure recombinant proteins in *E. coli* requires a successful combination of gene expression, protein solubility, and purification method.^[102] In general, efforts to improve the solubility of the protein have included the use of weak promoters, low cultivation temperature during protein productions, modified growth media, and fusion with solubility enhancing tags.^{[103],[104]}

Here, three ancestral reconstructions were recombinantly produced in *E. coli* BL21 (DE3). Under the T7 promoter, ANC1 was solubly overproduced best among all three recombinant enzymes. (Figure 17).

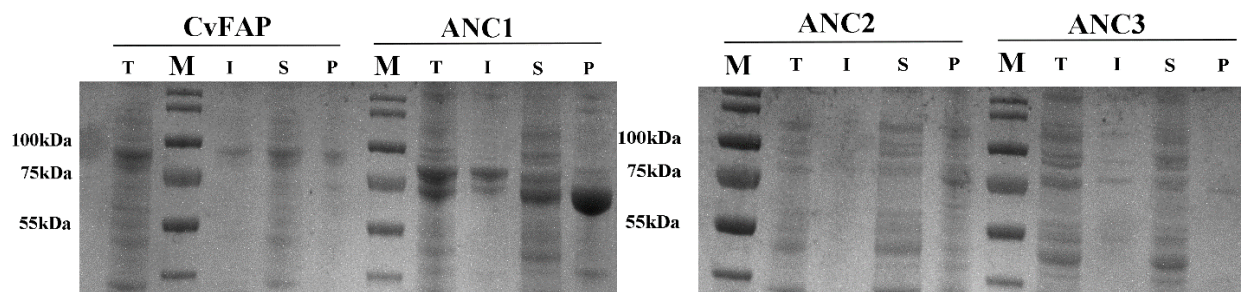


Figure 17. SDS-PAGE analysis of different fractions from Ni-affinity purification of CvFAP, ANC1, ANC2 and ANC3.

Lane M, molecular mass marker; Lane T, total cell lysate; Lane S, soluble fraction of cell lysate; Lane I, insoluble fraction; Lane P, purified enzyme. The molecular weight of the CvFAP: 77 kDa; ANC1: 65 kDa; ANC2: 65 kDa; ANC3: 65 kDa.

After one-step Ni-affinity purification, ANC1 showed a higher purification yield than CvFAP, about 12 folds (Table 2). The final yield of purified ANC1, from 1 liter of *E. coli* bacterial culture, 24 mg of the recombinant protein could be obtained. In contrast, ANC2 and ANC3 yielded similar low purification yields as the wild-type CvFAP.

Table 2. Purification yield of CvFAP and ancestor enzymes.

	CvFAP	ANC1	ANC2	ANC3
Purification Yield (mg/L)	2.34 ± 0.2	24.09 ± 1.0	1.4 ± 0.2	2.6 ± 0.7

After production, we verified whether these enzymes retained photodecarboxylation activity when palmitic acid was used as the substrate. Purified enzymes were used as catalysts for activity measurements. The results evidenced that ANC1 was the only active ancestor, ANC2 and ANC3 showed no photodecarboxylation activity. These experiments convincingly demonstrated that the quality of the phylogenetic tree and the evolutionary model selection is crucial for the reliability of reconstruction.

3.4. Study of FAPs expression level and activity by IPTG concentration or removing predicted chloroplast sequence

As already shown by several studies, the concentration of the inducer IPTG is crucial for soluble protein production.^{[105],[106]} Besides the expression level, the high amount of IPTG might increase non-functionally expressed protein levels.^{[107],[108]}

We could raise the soluble production of CvFAP by using different promoters and different concentrations of inducers from our previous study. A recently published study demonstrated that higher expression levels could be achieved by removing the predicted chloroplast sequence in CvFAP.^[57] Also, as Scrutton *et al.* demonstrated, the catalytic activity of CvFAP was affected by the plasmid backbone used and the position of fused

purification tags to CvFAP.^[59] Therefore, we attempted to further improve CvFAP and ANC1 by increasing IPTG concentration to 0.5 mM or removing the predicted chloroplast sequence in ANC1.

For the ANC1 study, we designed two variants based on the ANC1 sequence named sANC1 and TrxA-sANC1. The sequence of sANC1 lacks the predicted chloroplast target sequence at the N-terminus and comprises residues 62-623 of ANC1. The fusion soluble tag (TrxA) increased the soluble production of the enzyme. Another variant of sANC1 was designed, which fused the TrxA tag at the N-terminus of sANC1. As shown in Figure 18, both variants were soluble produced and could be purified through Ni-affinity chromatography.

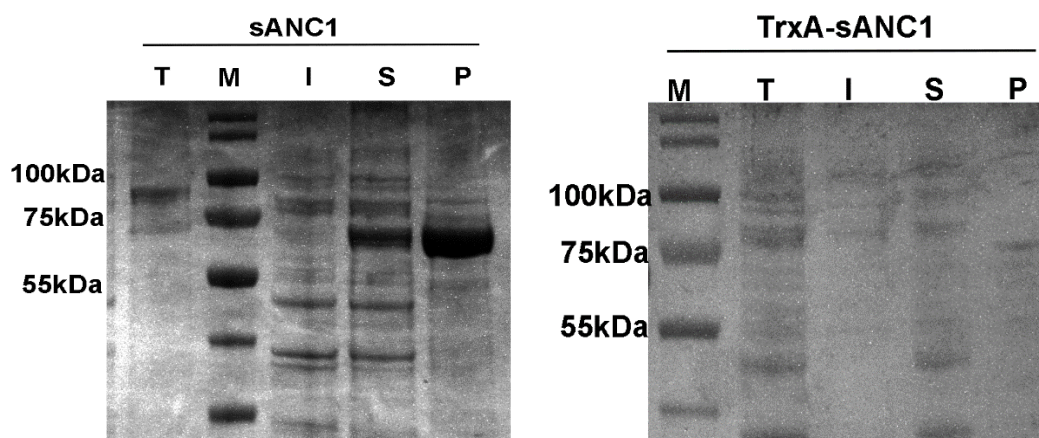


Figure 18. SDS-PAGE analysis of different fractions from Ni-affinity purification of sANC1 and TrxA-sANC1.

Lane M, molecular mass marker; Lane T, total cell lysate; Lane S, soluble fraction of cell lysate; Lane I, insoluble fraction; Lane P, purified enzyme. Molecular weight of the CvFAP: 77.0 kDa; ANC1: 65.0 kDa; sANC1: 60.0 kDa.

As shown in Table 3, for CvFAP, the purification yield for CvFAP was increased from 2 mg/L (with 0.1 mM IPTG) to 13 mg/L (with 0.5 mM IPTG), and 16 mg/L for the arabinose and the tetracycline controlled expression systems, respectively. Therefore, removing the predicted chloroplast sequence favored the improvement of CvFAP expression level. Stronger promoters or decreasing IPTG concentration enable to improve CvFAP expression level as well.

When expression was induced with different IPTG concentrations, the purification yield of ANC1 increased from 24 mg/L (with 0.1 mM IPTG) to 32 mg/L (with 0.5 mM IPTG). When removed the predicted chloroplast sequence of ANC1, the soluble production was 33 mg/L (0.5 mM IPTG). As we expected, an improvement of purification yield for ANC1 was observed when increasing IPTG concentration. Moreover, removing the chloroplast sequence did not improve the purification yield for ANC1. When the TrxA soluble tag was fused to the N-terminus of sANC1, the purification yield was 2.3 mg/L, which was much lower than for ANC1 and CvFAP purification yields with the same expression conditions.

Table 3. Purification yields of CvFAP, ANC1, and sANC1 enzyme.
(mg/ L cultivation volume)

Expression vector	Purification yield (mg/L)	Expression vector	Purification yield (mg/L)
pBAD-CvFAP (0.1% arabinose)	16.9	pASK-CvFAP (200 µg/L Anhydrotetracycline)	15.9
pET28a-CvFAP (0.1 mM IPTG)	2.3 ± 0.2	pET28a-CvFAP (0.5 mM IPTG)	13.3 ± 2.3
pET28a-ANC1 (0.1 mM IPTG)	24.1 ± 1.0	pET28a-ANC1 (0.5 mM IPTG)	32.2 ± 2.5
pET28a-TrxA-sANC1 (0.5 mM IPTG)	2.3	pET28a-sANC1 (0.5 mM IPTG)	33.5 ± 4.7

Surprisingly, a higher concentration of IPTG resulted in a significant improvement in CvFAP specific activity. However, the stability was extremely low (Figure 19). This resulted in a 30-fold improvement in the volumetric activity of cell-free extracts compared to expression with 0.1 mM IPTG. For ANC1, the volumetric activity almost did not change. It appears that ANC1 production is more consistent at different IPTG concentrations. In contrast, the functional production of CvFAP is favored at higher IPTG concentrations. This indicates losing of catalytic activity could be reduced by performing the *in vitro* purification under red light,^[56] and increasing inducer concentration also could favor functional production of CvFAP.

For ANC1, removing the predicted chloroplast target sequence suggested no significantly specific activity difference for C16. The fusion TrxA-sANC1 protein proved no activity, TrxA soluble tag was around 11 kDa, the misfolding of sANC1 caused the low soluble expression and losing activity.

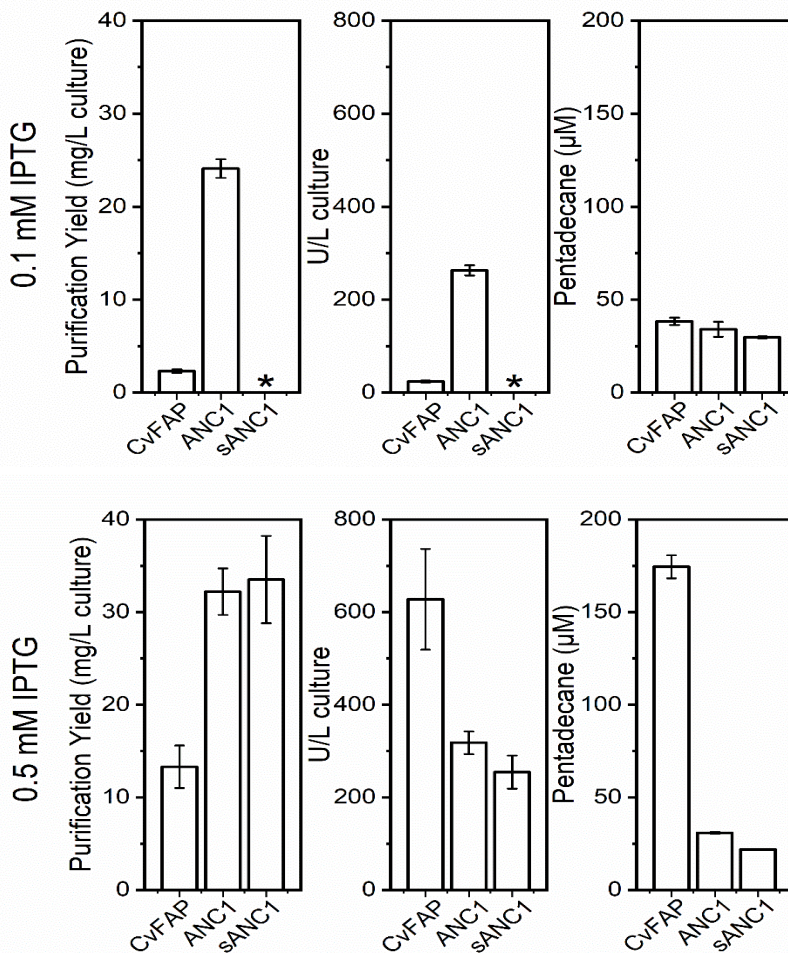


Figure 19. Comparison between CvFAP, ANC1, and sANC1 produced using two different IPTG concentrations.

The plots on the left report production yield as mg/L culture, the plots in the center report Units per liter culture, and the plots on the right show pentadecane production using palmitic acid as substrate after 1 h using 4 μM of purified enzymes. * Not biological triplicated group.

3.5. Photoinactivation of fatty acid photodecarboxylases

Since FAPs are the light-dependent enzyme compared to other proteins, the light attributed to catalytic instability was the subject for mechanistic investigations and exploitation of FAPs. Since CvFAP is less catalytically efficient after purification than in the crude cell-free extract, the loss of catalytic activity could be reduced by performing the *in vitro* purification under red light.^[56] One explanation for the photoinactivation was the light exposure synthesis protein-based organic radical accumulation, and the photoactivated FAP retains overall protein structure and bound FAD. These results supported our observation about the photoinactivation of FAPs (Figure 20). When FAPs were exposed under blue light ($\lambda=455$ nm) for a short period, the catalytic activity was reduced dramatically. When exposed to blue light ($\lambda=455$ nm) for 5 minutes, CvFAP and

sANC1 presented 2-fold higher residue activity than ANC1; after 15 minutes, the residue activity for CvFAP, sANC1, and ANC1 was close to each other.

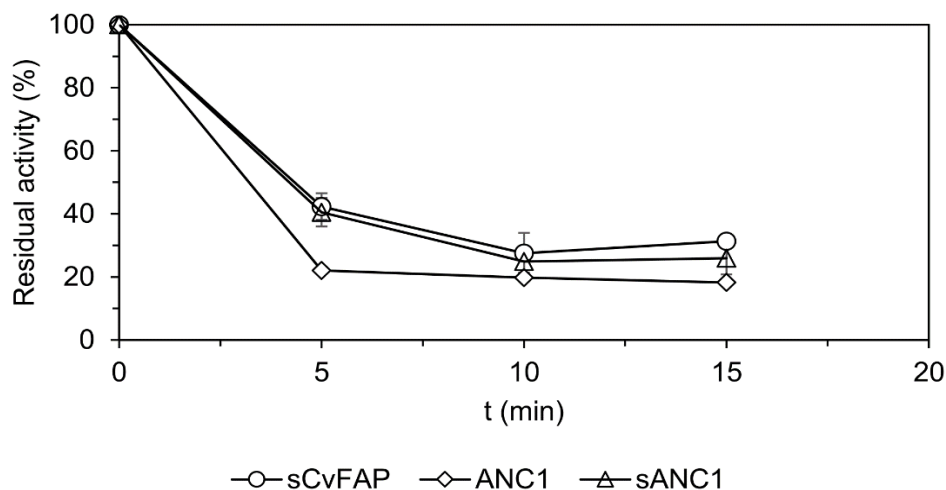


Figure 20. Comparison of sCvFAP, ANC1, and sANC1 residual activities after being exposed to light at 455 nm for the stated amount of time.

To confirm whether the photoinactivation was irreversible, free FAD was added to the reaction systems. No improvement of the catalytic activity was observed for FAPs, as visualized in Figure 21. This indicated that the poor catalytic stability of FAPs was attributed to light exposure, and the inactivation process was fast and irreversible. As demonstrated above, the higher concentration of IPTG favored the functional production of CvFAP, which might be due to the accelerated folding process reduced the light exposure period, which leads to the inactivation of FAPs.

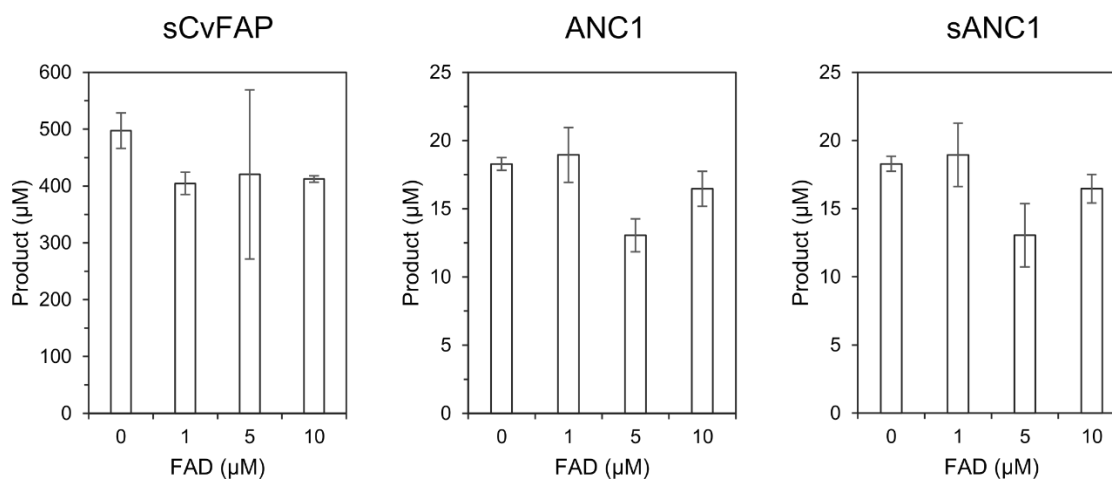


Figure 21. Comparison of sCvFAP, ANC1, and sANC1 product production at different external FAD concentrations.

In all reactions, the enzyme concentration was kept at 5 μM , and reactions were performed for 1 h.

We observed that the denaturation of CvFAP yielded pronounced decreasing fluorescent intensity, which was different from other flavoproteins. During the denaturation of flavoproteins, the secondary and tertiary structure of the protein is disrupted, and interactions with flavin break down, usually leading to significant increases in the flavin fluorescence.

From the observation of CvFAP, a blue shift in the FAD peak fluorescent feature (from 550 nm to 540 nm) and a decrease in FAD fluorescence emission after denaturation were observed. This indicates that light exposure altered the FAD cofactor's environment/electronic fluorescent properties in CvFAP (Figure 23).

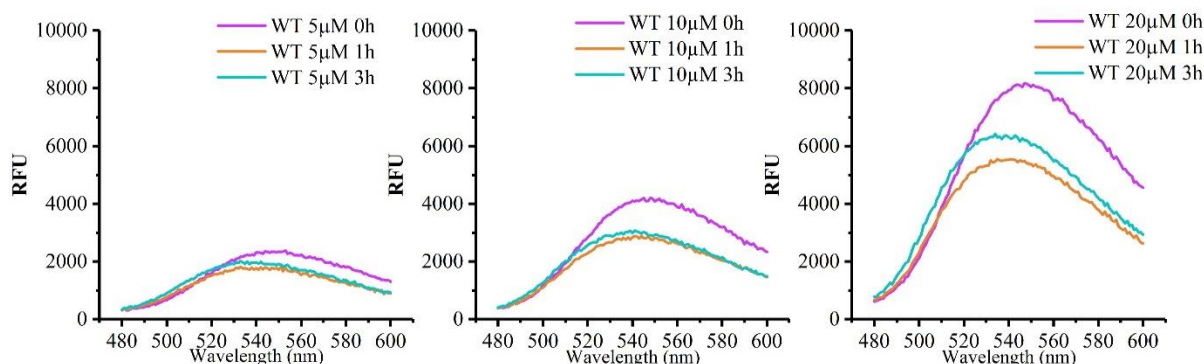


Figure 22. Time-resolved emission spectroscopy for different concentrations of CvFAP. Measurement condition: Purified CvFAP= 5 μ M, 10 μ M, 20 μ M; Tris-HCl buffer pH 8.5 (100 mM); 30°C. Fluorimeter setting was as follow: λ_{ex} =450 nm, λ_{em} =480-600 nm, Gain=100 nm. The emission maxima λ_{max} of CvFAP shifted from 550 nm to 540 nm after denaturation.

As the results showed in Figure 23 B, under the same measurement condition, the selected control flavin-containing monooxygenase CahJ with FAD cofactor showed increasing fluorescence intensity, expected from the unquenched FAD released from the denatured protein. The time-resolved emission spectroscopy of ANC1 (Figure 23 C), with the same concentration of purified enzyme, the fluorescence signal of ANC1 was much stronger than CvFAP. This difference in quantum yield between the two enzymes suggested that the isoalloxazine ring might be less rigidly fixed in ANC1. Also, there was no blue-shifted fluorescent peak for ANC1. With the same concentration, it also needs to be mentioned, with the ANC1 fluorescent signal much higher than free FAD under the same concentration. The fluorescence intensity difference between CvFAP and ANC1 indicated that protein structure plays an essential role in the FAD quenching/folding effect.

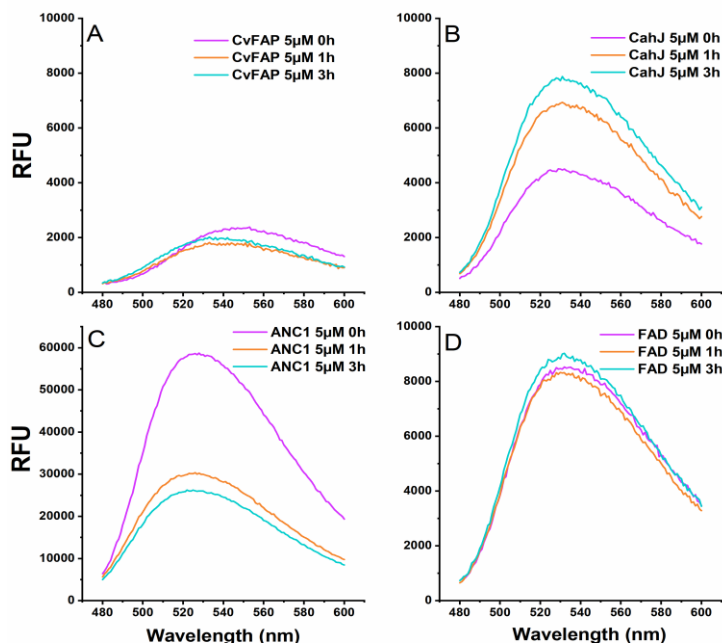


Figure 23. Time-resolved emission spectroscopy of CvFAP, ANC1.

Flavin-containing monooxygenase CahJ was the positive control, and FAD was the negative control. CvFAP and ANC1 showed an opposite trend with the FAD signal decreasing compared to the positive control; FAD remained relatively constant. All measurements were performed using 5 μM of purified enzyme in Tris-HCl buffer pH 8.5 (100 mM); 30°C. Fluorimeter setting are as follow: $\lambda_{\text{ex}}=450$ nm, $\lambda_{\text{em}}=480\text{-}600$ nm, Gain=100 nm.

The photoinactivation and difference fluorescence spectra between CvFAP and ANC1 we observed were interesting, since for FAPs mechanism study, how the residues in active site participated in adding proton/hydrogen atom and transfer electron remain to be clarified. Till now, only two strictly conserved residues Arg451^[54] and Cys432^[55] in FAP, were demonstrated might act as proton donors (Figure 24). Therefore, elucidating the differences of active site residues between CvFAP and ANC1, especially the amino acid residues near the FAD bonding site, might further understand the kinetic and chemical mechanisms of FAPs.

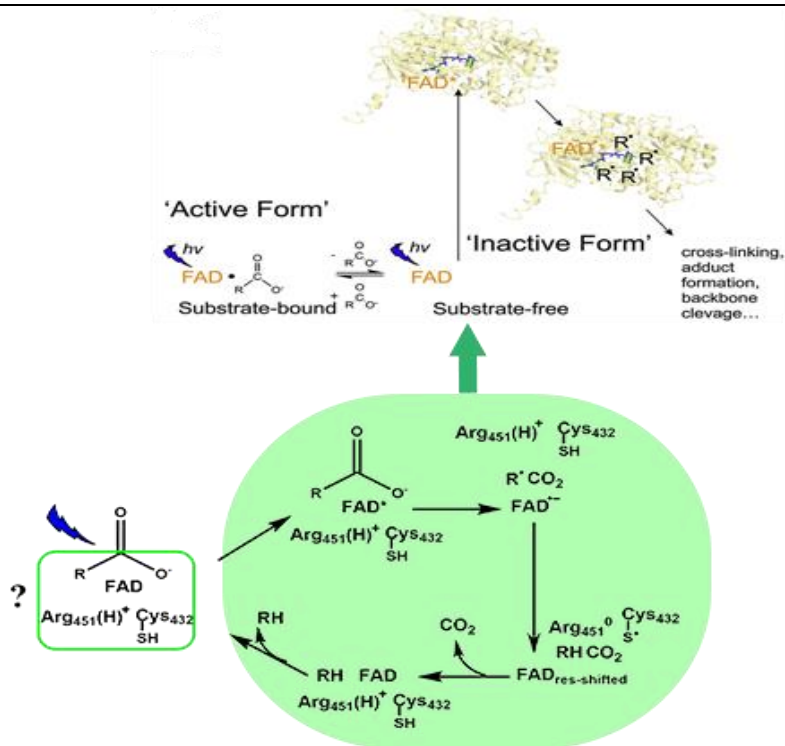


Figure 24. Proposed photocycle of CvFAP.

The green circle presented the proposed photoinactivation mechanism of CvFAP.

3.6. ANC1 and sANC1 have higher Tm and thermostability.

In many cases, ancestral proteins are more thermostable than modern proteins. This is often explained by a higher average earth temperature in prehistoric times.^[109] We anticipated that ANC1 would show higher thermostability compared to CvFAP. Therefore, we determined the melting temperatures (Tm) for CvFAP and ANC1 using the ThermoFAD method.^[110]

As presented above, the denaturation of CvFAP and ANC1 yielded a very pronounced decreasing fluorescence intensity. This contrasted with the known increasing fluorescence observed for a flavin-dependent monooxygenase CahJ used as control (Figure 23). Thus, in this case, the Tm was identified by plotting the negative of the first derivative of the fluorescence emission as a function of temperature ($-dF/dT$).

At first, the fluorescent signal of CvFAP was low and presented undefined peaks. To resolve this, we performed the thermoFAD experiment with CvFAP expressed under different promoters to obtain more defined melting temperature curves of CvFAP for better comparison with the profiles observed for ANC1 (Figure 25). A profile with two peaks in the derivative curve was found in all cases, indicating that two distinct domains unfold at different temperatures. This is in agreement with observations from *Sorigue et al.*^[111] who described a structure with two distinct domains, one for binding FAD and one for the substrate, respectively. We hypothesized that for the expression of the CvFAP gene under

the control of the T7 promoter, a mixed population of correctly folded and unfolded CvFAP is obtained that resulted in the undefined profile (Figure 25, C). Producing CvFAP from vectors with a weaker promoter gave similar production yields but resulted in a more homogeneous enzyme population with defined melting temperature profiles (Figure 25, A and B).

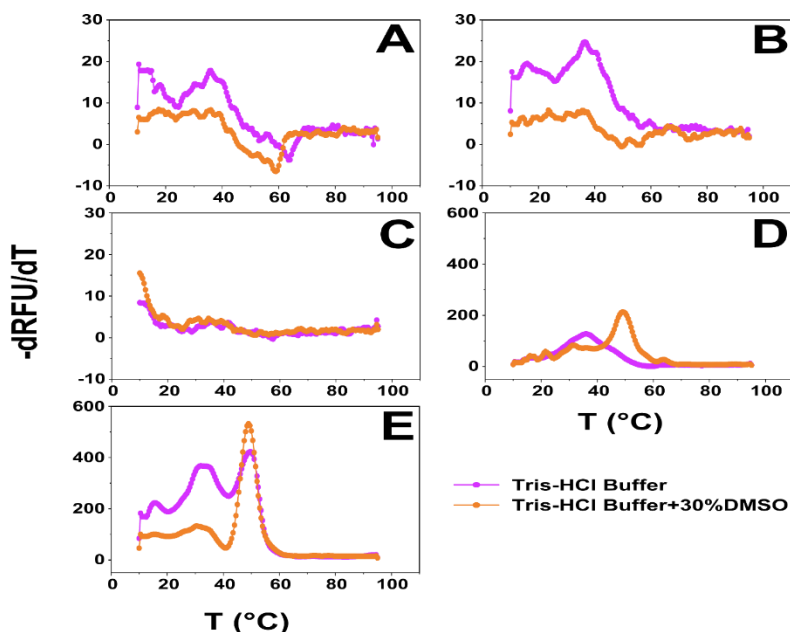


Figure 25. First derivative curves of the melting temperature curves are obtained in fluorescence measurements.

A: CvFAP produced in *E. coli* TOP10 (pBAD-CvFAP); B: CvFAP produced in *E. coli* BL21 (DE3) (pASK-CvFAP); C: CvFAP produced in *E. coli* BL21 (DE3) (pET28a-CvFAP); D: ANC1 produced in *E. coli* BL21 (DE3) (pET28a-ANC1); E: sANC1 produced in *E. coli* BL21 (DE3) (pET28a-sANC1).

Here, ANC1 and sANC1 showed a significantly higher melting temperature for both domains with an increase of 15 °C and 13 °C, respectively (Table 4). Furthermore, the addition of 30% DMSO (Figure 25, orange curves) seems to accelerate the denaturation of the second peak, as a faster denaturation results in a higher $-dF/dT$. Possibly, the second peak represents the substrate-binding domain, as it should be more hydrophobic to interact with fatty acids. Thus, while interacting more easily with DMSO, it unfolds at a faster rate.

Table 4. Comparison of T_m of both peaks determined by ThermoFAD for CvFAP, ANC1, and sANC1 in aqueous buffer with or without 30% DMSO.

	$T_m(^{\circ}\text{C})$	
	Tris-HCl Buffer	Tris-HCl Buffer+30%DMSO
pBAD-CvFAP	14/36	18/36
pASK-CvFAP	16/36.5	24/35.5
pET28a-CvFAP	n.d.	n.d.
pET28a-ANC1	36/44	31/49
pET28a-sANC1	36/49.5	31/49

n.d. not determined

To validate this data, we further determined the thermostability of ANC1, sANC1, and CvFAP by measuring the residual activity at 30 °C after incubating the purified enzymes at a range of temperatures (Figure 26). The plots confirm that ANC1 and sANC1 are more thermostable photodecarboxylases compared to CvFAP.

Again, different IPTG concentrations in protein production had a more significant effect on CvFAP than the resurrected enzymes. In particular, CvFAP produced with 0.5 mM IPTG seems to be approximately three times less thermostable compared to CvFAP produced with 0.1 mM IPTG. A higher concentration of IPTG yields more active enzymes far more prone to aggregation and thermal denaturation (Figure 26). The higher concentration of the inducer might accelerate the folding process of CvFAP. Thus, the light exposure period was reduced.

ANC1 and sANC1 seem to be more robust to changes in expression conditions and maintained similar residual activity at both IPTG concentrations. With 0.1 mM IPTG, ANC1 showed a 2-fold improvement than CvFAP in residual activity at 50 °C, whereas it displayed almost 20-fold improvement when produced at 0.5 mM IPTG at 50 °C. All in all, the higher thermostability determined with the ThermoFAD method could be validated with activity measurements confirming that ANC1 and sANC1 are significantly more thermostable than the wild-type enzyme.

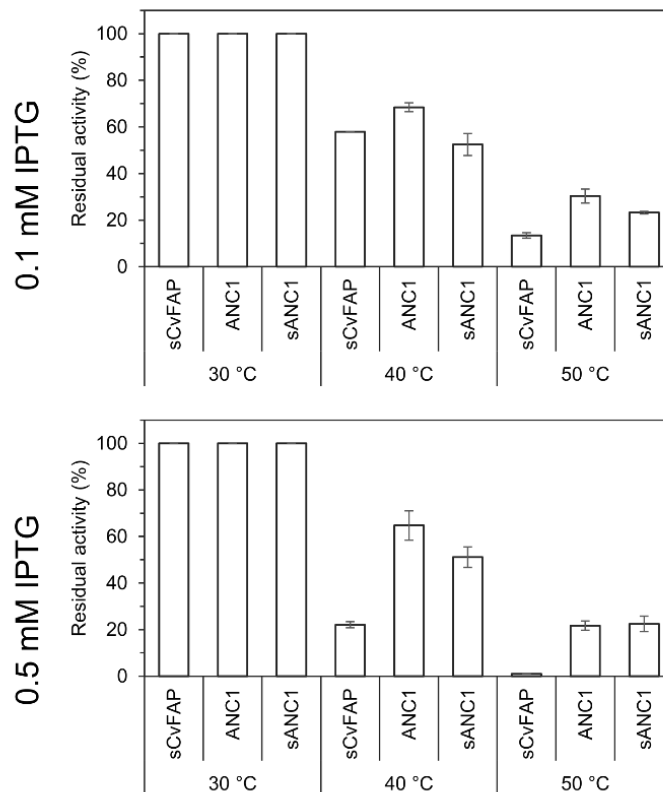


Figure 26. Thermostability of CvFAP, ANC1, and SANC1 produced using different IPTG concentrations (0.1 mM in A and 0.5 mM in B).

Purified enzymes were incubated at different temperatures for 10 min, and then the reactions were performed with gentle magnetic stirring at 30 °C in a total volume of 200 μ L Tris-HCl buffer (pH 8.5, 100 mM) containing 30% DMSO.

One proven method to increase enzyme stability is to increase the packing at the enzyme's core with hydrophobic interaction and increase the hydrophilicity of surface residues for better interaction with aqueous systems.^{[109], [112]} Thus, we investigated whether we could observe significant changes in surface residues. As anticipated, most of the substitutions on the outer surface of ANC1 were towards more hydrophilic amino acids (Figure 27). In contrast, we could not identify variants in the protein core that may explain the improvement in the production yield of ANC1 without compromising the functional activity.

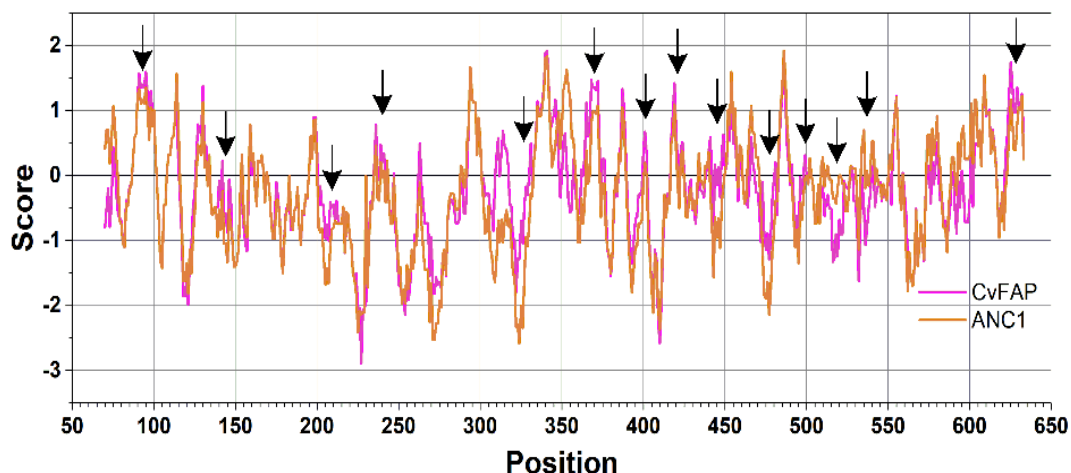


Figure 27. ProtScale hydrophobicity graph for CvFAP (pink) and ANC1 (orange). A positive score is associated with higher hydrophobicity. Arrows designate the surface residues where appreciable changes are visible.

As shown in Figure 28, since the ANC1 showed more thermostability compared to CvFAP, as anticipated, more rigid regions in ANC1 were identified comparing with CvFAP. The rigid region in ANC1 might interact with other amino acids tightly. The flexible regions in CvFAP are mainly located at loops on the surface and N-terminal of CvFAP. Therefore, the study of ANC1 and CvFAP may provide insights into identifying important sites of FAPs. Making changes with the flexible regions or before the flexible regions may further improve CvFAP stability.

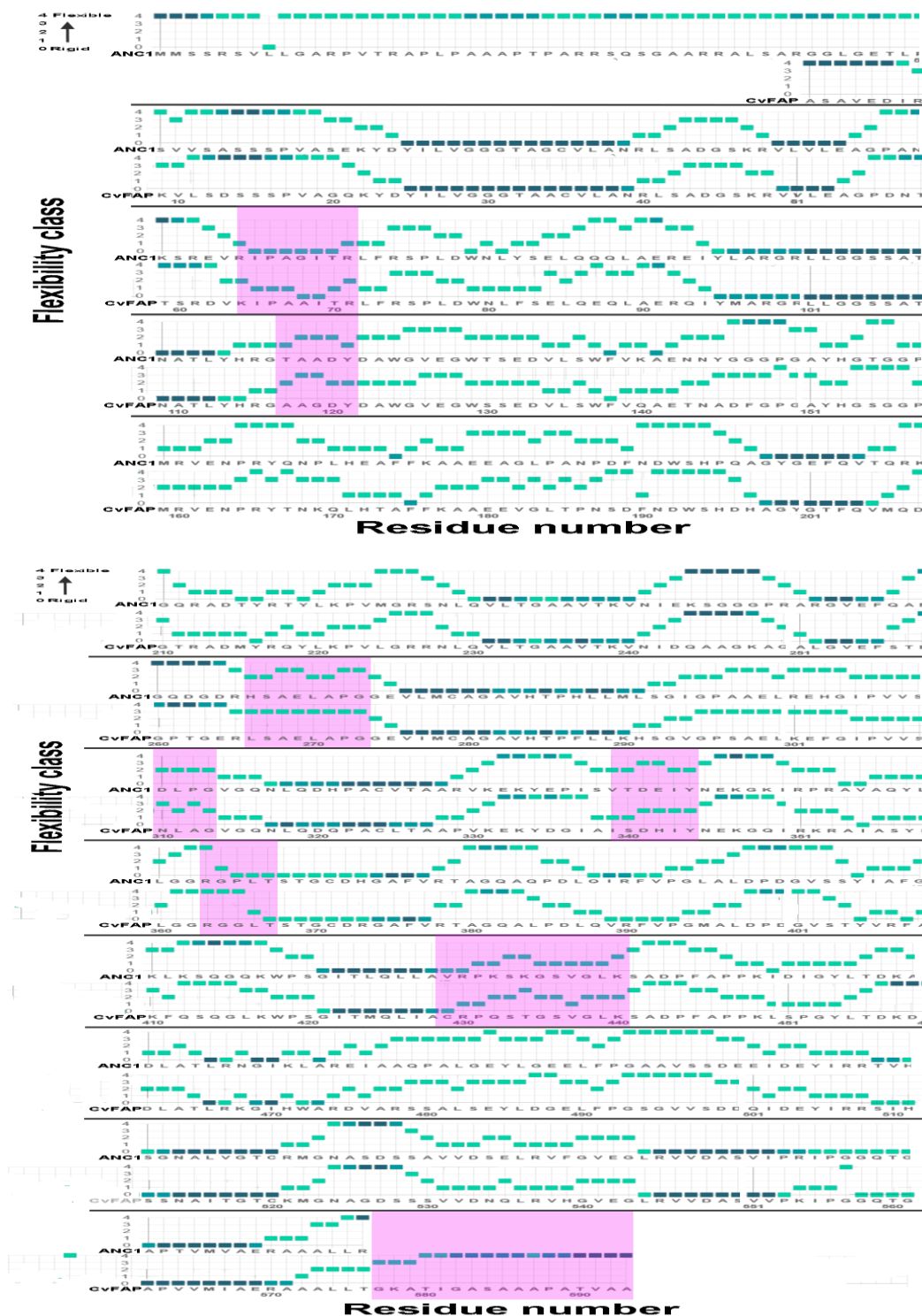


Figure 28. Flexibility comparison of CvFAP and ANC1. The different residues are highlighted in pink. The prediction was made via MEDUSA.^[113]

3.7. Preparative-scale production of pentadecane

The sustainability of biocatalyst is essential, particularly for industrial biocatalysis targeting fuel and chemical production. As shown above, ANC1, compared with CvFAP, exhibited higher stability and soluble production but relatively low activity. Thus, we prompted the pentadecane production by opting for lyophilized whole cells as catalysts. Since the reduced catalytic activity of FAPs is attributed to light exposure during expression and purification processes, the whole cell as the catalyst is optimal for maintaining enzyme activity. In our case, since the fatty acid was substrates, the higher acid concentration leads to the significant drop of pH,^[114] the residual cell wall compounds could act as a protective manner for the intracellular enzyme, thus perform enzymatic reactions under harsh conditions.^[115]

As shown in Figure 29, substrate (palmitic acid) concentration was increased to 100 mM. After the reaction was performed for 9 hours, CvFAP produced twice higher pentadecane (43 mM) than ANC1 (25 mM). However, up to the 20 h time point of reaction, the ANC1-mediated reaction proceeded at a constant reaction rate, the same as the first 9 hours. In contrast, the reaction rate of CvFAP was decreased. After 30 hours of reaction, pentadecane's final production was similar between CvFAP (62 mM) and ANC1 (52 mM). Consequently, ANC1 is more stable and sustainable for harsher reaction conditions. Still, a major limiting factor for ANC1 is the low catalytic activity compared with CvFAP.

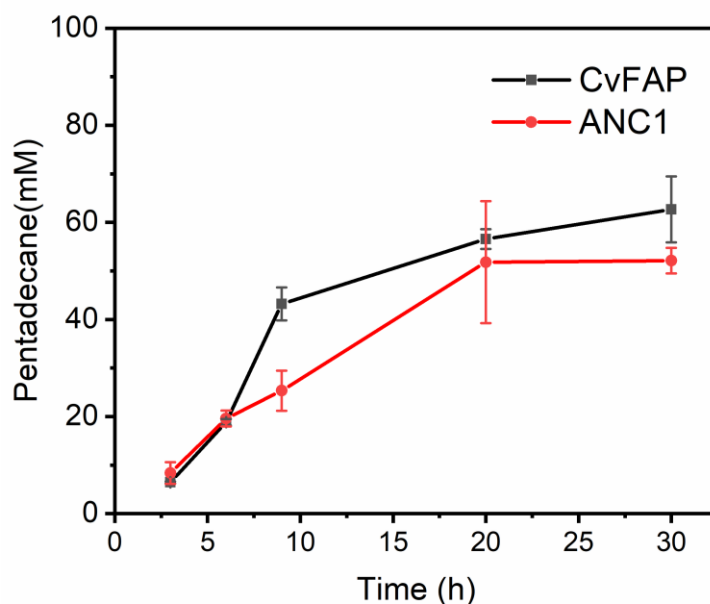


Figure 29. Time course of pentadecane production of lyophilized whole *E.coli* BL21(DE3) expressing CvFAP or ANC1.

The reaction was performed with gentle magnetic stirring at 30 °C in a total volume of 1.0 mL Tris-HCl buffer (pH 8.5, 100 mM) containing 30% DMSO. The reaction conditions: [Palmitic acid] =100 mM, [lyophilized whole cell] =10 mg.

3.8. Influence of the putative chloroplast target sequence on the selectivity of ANC1.

As reported earlier, removing the predicted chloroplast-target sequence in CvFAP increased the solubility and production yield significantly and could slightly improve catalytic activity.^[57] Since the ASR already resulted in a shorter N-terminus than the wild type, as shown in our previous study (Figure 17), removing the predicted chloroplast target sequence in ANC1 showed no significant improvement for soluble production and activity.

CvFAP showed different substrate preferences for oleic acid, linoleic acid, and saturated counterpart stearic acid as a substrate in previous studies.^[65] Therefore, we selected these three substrates for ANC1 and sANC1 to investigate if the resurrected enzymes display a different preference than the wild type. Interestingly, ANC1 showed no activity for the unsaturated substrates whereas, sANC1 showed conversion for all three but with a high preference for the unsaturated counterpart. It appears that the chloroplast sequence influences ANC1 substrate preference as increasing the amount of enzyme did not show an increased conversion for the unsaturated substrate in ANC1 (data not shown). Furthermore, ANC1 and sANC1 being more thermostable, are also less flexible; thus, we believe that unsaturated fatty acids are less likely to fit in the active tunnel as they are also less flexible than their saturated counterparts. This might contribute to the higher ANC1 chemoselectivity towards unsaturated fatty acids compared to wild-type.

Table 5. Decarboxylation of ANC1 and sANC1 of various fatty acid substrates compared to wild-type CvFAP.

Fatty acid	Product (μM)		
	CvFAP	ANC1	sANC1
C18:0	1146.9	128.1	278.8
C18:1 Δ9	651.2	0	25.9
C18:2 Δ9,12	29.6	0	16.5

3.9. Applicability of FAPs in deep eutectic solvents

For large-scale applications of FAPs, the poor solubility of free fatty acids in water limits the load of the substrate, and furthermore, a co-solvent is always needed to further increase the substrate solubility. Organic solvents can interact with the hydrophobic patches of proteins perturbing their folding/unfolding equilibrium in aqueous solutions.^[85] Deep Eutectic Solvent (DESs) have often been used as an alternative to organic solvents.^{[88],[89]} DES and DES/water mixtures have been shown to partially preserve enzyme activity leading to improved performances compared to similar setups using organic solvents.^[116] Several attempts were also made in biocatalysis where DES has

been used as (co)-solvents resulting in increased substrate solubility and enzyme stability.^{[91],[92]}

In our case, the DESs are more desirable as co-solvent for photodecarboxylase. On one side, in photodecarboxylation reaction, the main side product is CO₂, but DESs enable to absorb the CO₂,^[117] which means DESs may facilitate alkane production. On the other side, diverse starting compounds could prepare hydrophobic DESs, ranging from long-chain quaternary ammonium halides to natural hydrophobic compounds, such as menthol and diverse fatty acids.^{[118]–[120]} That provides the possibility through the hydrophobic fatty acid-based DESs, to overcome the typical bottleneck in decarboxylation, which are the poor solubility of fatty acid and challenge of loading substrates.

Initially, two hydrophilic choline chloride (ChCl) based DES (ChCl: Glycerol 1:2, ChCl: Urea 1:2) were tested in FAPs catalyzed decarboxylation of fatty acids. The DESs were mixed with an aqueous buffer in different volume ratios to test the tolerance of FAPs for DESs.

Both CvFAP and ANC1 catalyzed the decarboxylation in ChCl: Glycerol DES, but only CvFAP showed activity in ChCl: Urea. In ChCl: Glycerol DES, no considerable improvement was observed for CvFAP compared to the control reaction system with 30% DMSO as co-solvent. On the contrary, even the activity of CvFAP in ChCl: Gly DES decreased dramatically, ANC1 appeared 9-fold higher activity in ChCl: Glycerol DES than in aqueous solution.

Table 6. Conversion of palmitic acid to pentadecane using different mixtures of DES and water ^a

Entry	Reaction medium (DES/H ₂ O) ^c	Conversion(%) ^b					
		ChCl: Gly ^d		ChCl:Urea ^d		30%DMSO	
		CvFAP	ANC1	CvFAP	ANC1	CvFAP	ANC1
1	0					23.6	1.6
2	50%	15.7	14.9	4.6	0		
3	60%	8.7	5.8	2.7	0		
4	70%	8.1	4.2	1.4	0		
5	80%	6.7	ND	1.8	ND		
6	90%	1.4	ND	1.4	ND		

^a Reaction conditions: The photoenzymatic decarboxylation reaction was performed in the respective reaction medium at the given ratio. With gentle magnetic stirring at 30 °C in a total volume of 1.0 mL Tris-HCl buffer (pH 8.5, 250 mM) containing 30% DMSO or DES as co-solvents. The final conditions: [Palmitic acid] =10 mM, [Purified enzymes] =10 μM, [Reaction time] = 20 h.

^b Determined by GC-FID. ND: Not Detected.

^c The DES was mixed with water in volume ratio. For example: 50%= 50%DES: 50% H₂O.

^d The DES ChCl: Gly (Gly=Glycerol) in a 1:2 molar ratio; ChCl: Urea in a 1:2 molar ratio.

Next, we attempt to further investigate individual enzyme behaviors by determining the melting temperature profiles of enzymes in the desired buffer. Considering ChCl, Glycerol and Urea could be the potent denaturants in the reaction medium, the thermal denaturation curves with different concentrations (as displayed in Table 7) of ChCl, Glycerol, and Urea were determined, respectively.

Table 7. Ingredients concentration of DESs

DES	ChCl(mM)	Glycerol(mM)	Urea(mM)
50% ChCl:Glycerol	0.89	1.81	-
60% ChCl:Glycerol	1.28	2.61	-
70% ChCl:Glycerol	1.75	3.56	-
80% ChCl:Glycerol	2.29	4.65	-
90% ChCl:Glycerol	2.90	5.90	-
50% ChCl:Urea	1.15	-	2.29
60% ChCl:Urea	1.65	-	3.29
70% ChCl:Urea	2.24	-	4.47
80% ChCl:Urea	2.92	-	5.84
90% ChCl:Urea	3.69	-	7.38

The melting temperature profiles of CvFAP and ANC1 are displayed in Figure 30. For ANC1, two distinct domains unfolding signal were observed, but only one peak be observed from ChCl: Gly DES, the T_m increased ~10 °C. Glycerol as one ingredient may inhibit protein unfolding. The additive of glycerol in an aqueous buffer showed no influence for inhibiting protein unfolding. This indicated when the reactions were performed in ChCl: Glycerol DES mixture, the higher pentadecane production of ANC1 may be attributed to the DES mixture to maintain better protein stability. On the contrary, the mixture of DES showed no significant improvement for alkene production of CvFAP. For melting temperature profiles, as we observed before, in aqueous buffer, there were no defined melting peaks for CvFAP, but in ChCl: Glycerol DES hybrid system, one relatively defined T_m was observed compare with the system without DES, this also indicated ChCl: Glycerol DES could maintain stability for CvFAP, even not significantly compare with ANC1.

Table 8. Comparison of T_m of both peaks determined by ThermoFAD for CvFAP, ANC1, in Tris-HCl buffer with different DES components.

	$T_m(^{\circ}\text{C})$	
	ANC1	CvFAP
Tris-HCl Buffer	36/44	n.d.
Tris-HCl Buffer+30% DMSO	31/49	n.d.
Tris-HCl Buffer+50% ChCl:Glycerol	n.d. /60	n.d./46
Tris-HCl Buffer+70% ChCl:Glycerol	n.d. /65	n.d./51
Tris-HCl Buffer+90% ChCl:Glycerol	n.d. /65	n.d./51
Tris-HCl Buffer+50% ChCl:Urea	n.d. /45	n.d.
Tris-HCl Buffer+70% ChCl:Urea	n.d. /50	n.d.
Tris-HCl Buffer+90% ChCl:Urea	n.d. /49	n.d.
Tris-HCl Buffer+0.5 M ChCl	37/48	n.d.
Tris-HCl Buffer+0.9 M ChCl	n.d. /48	n.d.
Tris-HCl Buffer+1 M Glycerol	37/49	n.d.
Tris-HCl Buffer+1.8 M Glycerol	37/51	n.d.
Tris-HCl Buffer+1 M Urea	n.d. /41	n.d.
Tris-HCl Buffer+1.8 M Urea	n.d. /37	n.d.

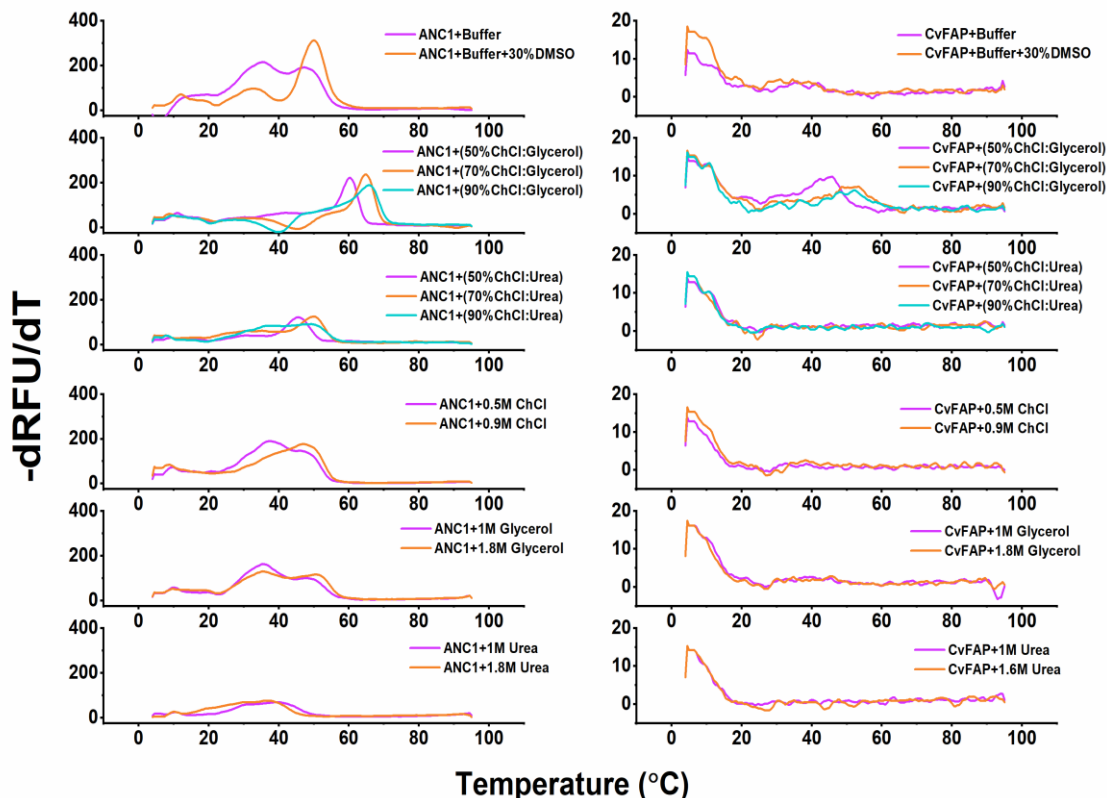


Figure 30. First derivative curves of the melting temperature curves were obtained in fluorescence measurements for CvFAP and ANC1 in different DESs concentrations. DESs components were tested separately as control. 5 μ M purified protein was added together with the buffers and components into the wells of the RT-PCR instrument. Next, a temperature gradient is applied, starting from 4 $^{\circ}$ C and increasing to 95 $^{\circ}$ C, measuring the fluorescence signal every 0.5 $^{\circ}$ C, every temperature keep for 10 s. RT-PCR instrument for ThermoFAD is CFX96 TouchTM, the detector filter fluorescence emission of 510-530 nm.

We envisioned tackling the solubility hurdles of palmitic acid in aqueous systems by using DESs, which are known to have solubilizing properties for some low solubility substrates. However, the hydrophilic DESs (ChCl: Glycerol and ChCl: Urea) we selected could not enhance the solubility of fatty acid in the reaction system. When loaded 100 mM of palmitic acid in DMSO aqueous and hydrophilic DESs systems, both CvFAP and ANC1 showed lower production of pentadecane. We observed that palmitic acid was less soluble in the DESs mixture than in the DMSO aqueous system. In the DESs system, the palmitic acid is saturated and partially precipitated.

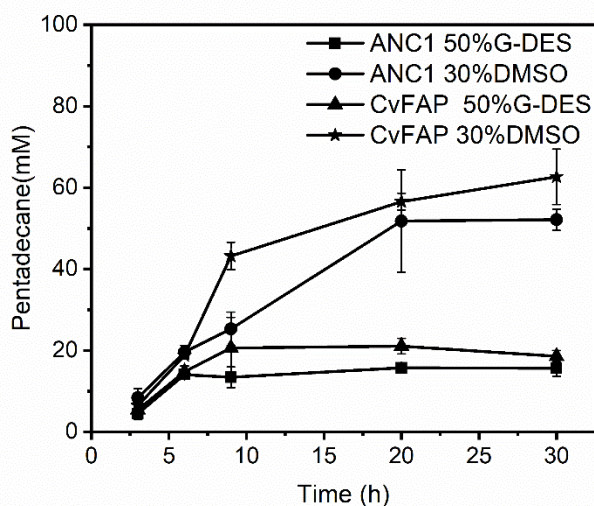


Figure 31. Time course of pentadecane production of CvFAP and ANC1 in DESs.

We also attempted to optimize the DESs systems by solving palmitic acid into DMSO firstly then mixing with DES. Yet, no production of pentadecane being observed in this waterless reaction due to the low pH of the DESs (for both ChCl: Glycerol and ChCl: Urea system, pH= 6) was harmful to the enzyme. As reported before, the activity of CvFAP was strictly dependent on pH.^[57]

Besides the hydrophilic DESs, we performed the reactions in the hydrophobic DESs, which are generated by fatty acid and quaternary ammonium salt.^[87] The advantage of hydrophobic DESs, fatty acids is the hydrogen bond donor and the substrates for FAPs.



Figure 32. Hydrophobic DES synthesized in this study

The quaternary ammonium salt tetrabutyl ammonium chloride ($N_{4444}\text{-Cl}$)^[121] was chosen as a hydrogen-bond acceptor, and we obtained the DES Palmitic acid: $N_{4444}\text{-Cl}$ in a molar ratio 2:1. However, after this DES was mixed with water, the fatty acid precipitated immediately. The phenomenon occurred due to the high solubility of $N_{4444}\text{-Cl}$ in water. The equilibrium between $N_{4444}\text{-Cl}$ and fatty acid was broken. Therefore, performing the reaction in waterless conditions might avoid precipitation of the fatty acid, the lyophilized whole cell was used as the catalyst. Nevertheless, there was no production of pentadecane being observed in this waterless reaction due to the extremely low pH (around 2) of the DES was harmful to the enzyme.

Table 9. Synthesis of Hydrophobic DES synthesized in this study

HBA	HBD	Molar ratio	HBA	HBD	Molar ratio
DL-menthol	C ₈	1:1	N ₄₄₄₄ -Cl	C ₈	-
DL-menthol	C ₁₀	1:1	N ₄₄₄₄ -Cl	C ₁₀	-
DL-menthol	C ₁₂	-	N ₄₄₄₄ -Cl	C ₁₂	1:1/1:2
DL-menthol	C ₁₄	-	N ₄₄₄₄ -Cl	C ₁₄	1:1/1:2
DL-menthol	C ₁₆	-	N ₄₄₄₄ -Cl	C ₁₆	2:1
DL-menthol	C ₁₈	1:1	N ₄₄₄₄ -Cl	C ₁₈	-

*HBA: Hydrogen Bond Acceptor; HBD: Hydrogen Bond Donor; N₄₄₄₄-Cl = tetrabutylammonium chloride.

As results showed in Table 10, we synthesized two new kinds of DES, caprylic acid (C₈) and capric acid (C₁₀), as hydrogen bound donor, DL-menthol was hydrogen acceptor. As presented in Table 10, no product was detected in the reaction that used DMSO as a co-solvent. Meanwhile, the reaction performed in C₈: DL-menthol mixtures system produced 1.3 mM Heptane. With C₁₀: DL-menthol DES, the production of Nonane was 0.7 mM, which was 3 times higher than in the reaction system with DMSO as co-solvent. As reported before, CvFAP showed a marked preference for long-chain fatty acid (C₁₆-C₁₇), the catalytic activity of CvFAP drop dramatically with a shorter carboxylic acid. As Hollmann's group reported, using the decoy molecular to fill up the vacant substrate access channel of CvFAP enables the conversion of various kinds of carboxylic acid.^[122] Our results demonstrated that the decarboxylation activity for FAPs could be maintained and improved in the DES reaction system. This paves the way for practical applications of FAPs to produce a high yield of products with a simple process.

In the DES mixtures, the fatty acids are not just the hydrogen-bound donor in the DESs, also the substrates for reaction. This provides the possibility to overcome the limitation of loading substrates and the setup of continuous processes, which leads to more suitable applications for sustainable chemistry.

Table 10. Production of Heptane and Nonane in DES

	Reaction medium	Products	CvFAP (mM)	ANC1 (mM)
1	C ₈ : DL-menthol	Heptane	1.32±0.2	1.36±0.2
2	C ₈ 30% DMSO	Heptane	0	0
3	C ₁₀ : DL-menthol	Nonane	0.76±0.18	0.71±0.25
4	C ₁₀ 30% DMSO	Nonane	0.21±0.03	0.21±0.03

^a Reaction conditions: The photoenzymatic decarboxylation reaction was performed in the respective reaction medium at the given ratio. With gentle magnetic stirring at 30 °C in a total volume of 1.0 mL Tris-HCl buffer (pH 8.5, 100 mM) containing 30% DMSO or DES as co-solvents. The final conditions: [Fatty acid] =100 mM, [lyophilized whole cell] =10 mg, [Reaction time] = 20 h.

3.10. Protein engineering of CvFAP stability and sANC1 catalytic activity

As our study before, through reconstructing ancestral genes, we obtained more soluble and stable ANC1 but lower catalytic activity compared with CvFAP. For obtaining more functional FAPs for applications, two options could be optimized: 1, Improving the stability of CvFAP; 2, Improving the activity of ANC1.

We applied the Rosseta software suite to identify mutations that were likely to stabilize the native form of CvFAP (PDB code 5NCC). With Rosseta prediction, all possible single-site mutations were introduced *in silico* in CvFAP structure (amounting to 11,560 mutations), and their changes in folding free energy ($\Delta\Delta G$) were computed. However, the size of the mutant library was still quite large. To increase the possibility of screening 'hits' and minimize the size of the library, the prediction from Hotspot Wizard was taken into consideration. We did not consider mutations of residues that interact with the catalytic triad to avoid alterations to the activity of the CvFAP.

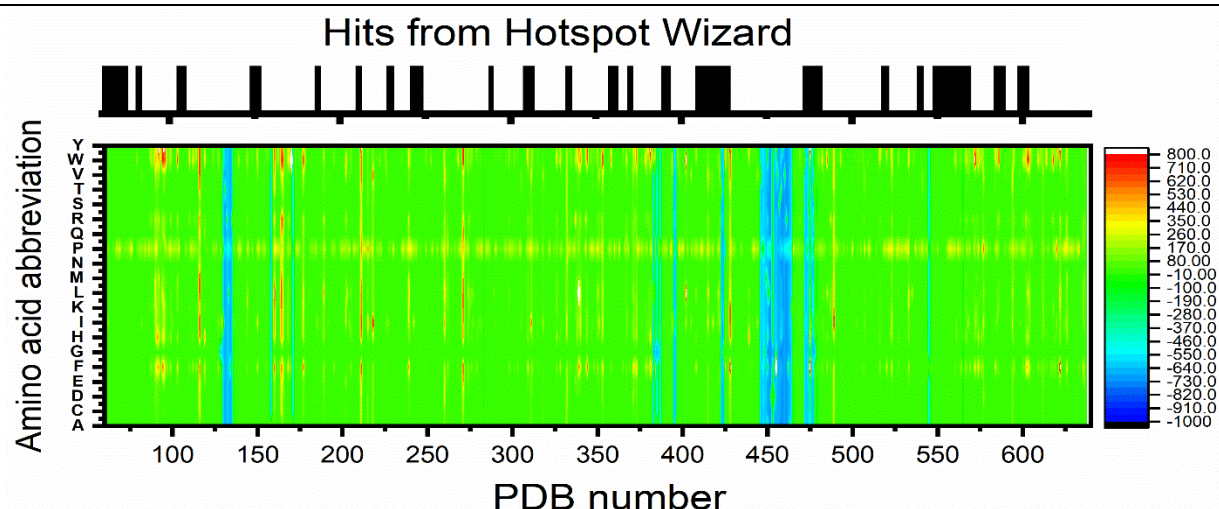


Figure 33. The stable hits from Hotspot Wizard prediction (Top); Point mutant (pmut) scanning score made via Rosetta scoring.

The more stable mutant presents a more negative score in Rosetta scoring. The stable hits from the Hotspot Wizard prediction presented in the black bar.

Based on the prediction, the first 12 stabilizing mutations were selected as listed in Table 11. Selected CvFAP variants were generated by site-directed mutagenesis *via* QuikChange protocol, and the desired point mutations were introduced in the CvFAP gene using pET28a-CvFAP as a template. We designed eight variants with altering amino acids at three different positions (E152, S473, E550), all three mutated residues located at CvFAP surface loops. Among these mutations, we obtained two mutated residues at E152, one mutated residue at S473, and five mutated residues at E550.

Table 11. Summary of predicted CvFAP variants

Variant	Mutations	Variant	Mutations
1	A62G	7	Q474N*
2	E152A*	8	E550R
3	E152N	9	E550N
4	E152G	10	E550G
5	E152S*	11	E550K
6	S473G	12	E550S

*not obtained

After confirming the success of obtaining desired mutations by Sanger sequencing, all eight mutation-constructed plasmids were transformed into *E.coli* BL21(DE3). The

expression level was analyzed from the SDS-PAGE gel (**Figure 34**). Eight mutations were successfully soluble expressed.

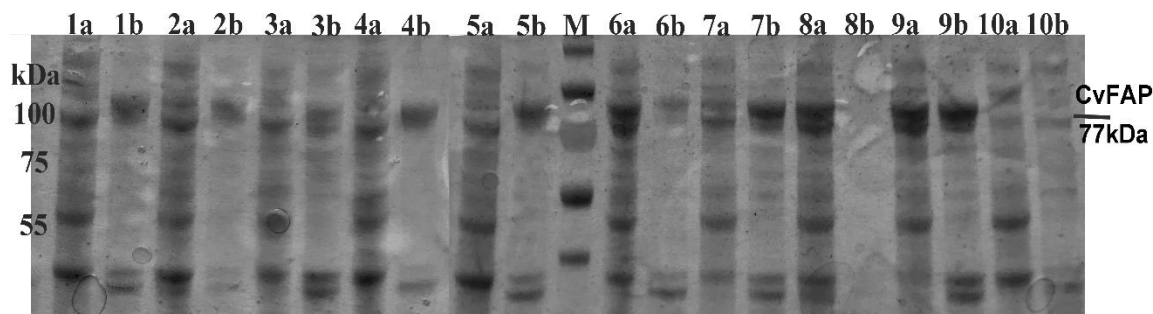


Figure 34. SDS-PAGE analysis of CvFAP mutants.

Lane 1a, insoluble fraction of CvFAP E550K, lane 1b soluble fraction of E550K; Lane 2a, insoluble fraction of CvFAP E550G, lane 2b soluble fraction of E550G; Lane 3a, insoluble fraction of CvFAP E550N, lane 3b soluble fraction of E550N; Lane 4a, insoluble fraction of CvFAP E550S, lane 4b soluble fraction of E550S; Lane 5a, insoluble fraction of CvFAP E550R, lane 5b soluble fraction of E550R; Lane 6a, insoluble fraction of CvFAP S473G, lane 6b soluble fraction of S473G; Lane 7a, insoluble fraction of CvFAP E152N, lane 7b soluble fraction of E152N; Lane 8a, insoluble fraction of CvFAP E152G, lane 8b soluble fraction of E152G; Lane 9a, insoluble fraction of CvFAP WT, lane 9b soluble fraction of WT; Lane 10a, insoluble fraction of pET28a, lane 10b soluble fraction of pET28a; The molecular weight of the CvFAP construct (6xHis-TrxA-TEV-CvFAP): 77.0 kDa. PageRuler Prestained Protein Ladder (Thermo Scientific) was used as standard.

Subsequently, the catalytic activity and thermostability of the eight mutants were experimentally determined (

Figure 35). Among the eight mutants, E152N, E152G, and S473G showed the same catalytic activity as wild-type CvFAP when the reaction was performed at 30 °C. After incubation at 40 °C for 30 min, only mutant S473G exhibited higher thermal residual activity than CvFAP. Notably, five mutants of E550 exhibited no activity, which indicated that this region is essential for catalytic activity.

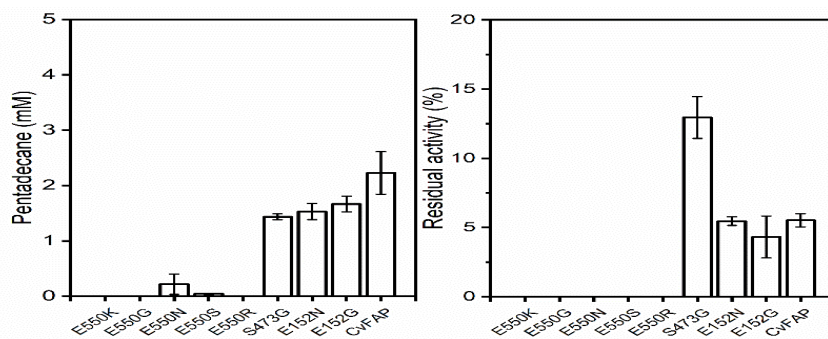


Figure 35. Production of pentadecane with CvFAP and its mutants (left) and thermostability of CvFAP and mutants (right).

The reaction was performed with gentle magnetic stirring at 30 °C in a total volume of 200 μ L Tris-HCl buffer (pH 8.5, 100 mM) containing 30% DMSO. The reaction conditions: [Palmitic acid] = 5 mM, [cell free extract] = 5 mg.

In the ThermoFAD method for determining melting temperature, as present in Figure 36, the variants E152G and S473G showed extended lifetime in the time-dependent inactivation experiments. The T_m for CvFAP was 45 °C; for E152G, T_m increased 5 °C. The variant S473G showed higher thermostability than CvFAP; as anticipated, the melting temperature was 55 °C, increased 10 °C comparing with CvFAP.

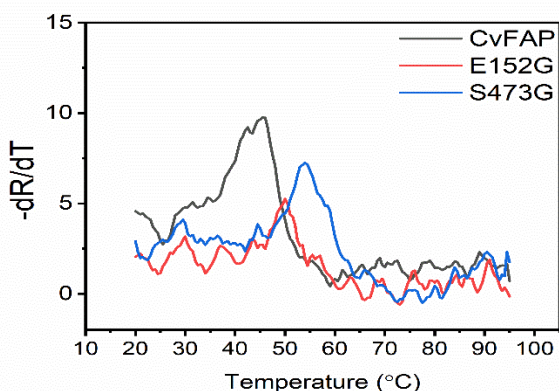


Figure 36. ThermoFAD experiment for CvFAP WT and mutant G152N, S473G.

The curves were obtained from thermal denaturation. 5 μ M purified protein was added together with the buffers into the wells of the RT-PCR instrument. Next, a temperature gradient is applied, starting from 4 °C and increasing to 95 °C, measuring the fluorescence signal every 0.5 °C, every temperature keep for 10 s. RT-PCR instrument for ThermoFAD is CFX96 TouchTM, the detector filter fluorescence emission of 510-530 nm.

From this part study, E550 was identified as an essential position related to the activity. After substituted at E550 with different substitutions, the enzymatic activity was abolished entirely. The investigation around E550 may provide more information about the structure and activity relationship. Until now, we only obtained eight variants, and two 'hit' residues influenced the stability of CvFAP. This indicated further study of the rest variants in the library would bring more information and improve CvFAP stability.

Since the sANC1 we obtained exhibited more soluble expression, higher folding efficiency, and stability. The sequence identity between CvFAP and sANC1 is 76%, the structure-based analysis can be used to identify the different regions between CvFAP and ANC1. The crystal structure of CvFAP (PDB: 5NCC) was selected as a modeling template, the structure visualized with PyMOL.

A study by Qi et al. identified the residues close to active sites influenced the activity and stereoselectivity of CvFAP.^[64] We compared the residues nearby the active site between ANC1 and CvFAP. The residues within 8 Å for active sites were selected. We found three residue differences between sANC1 and CvFAP. Further insight for difference details is shown in Figure 38. The residues in CvFAP, Leu386, Ser574, and Ser267 are different

from the residues locate at the same position for sANC1, and the residues were Val366, Gly552, and Thr247. We targeted to substitute V366 for V366L, G552 for G552S, and T247 for T247S of sANC1.

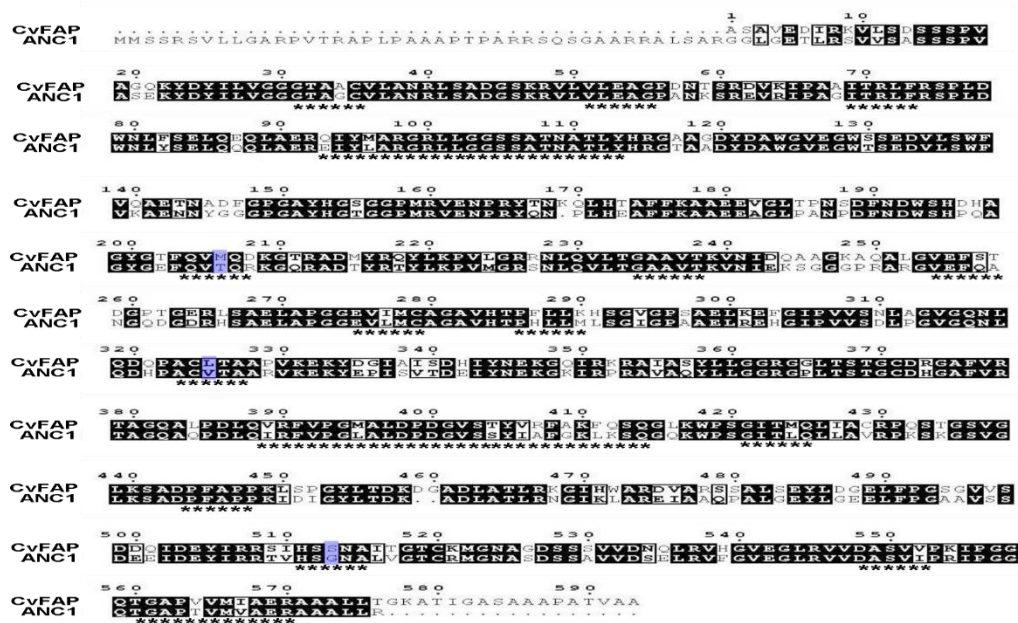


Figure 37. Sequence alignment of CvFAP and ANC1.

The alignment created with Clustal Omega^[123] and visualized with SnapeGene. The identical residues are highlighted in black. The residues within 10 Å for the activity site were labeled with the asterisk. The selected residues for the ANC1 activity study are highlighted in blue.

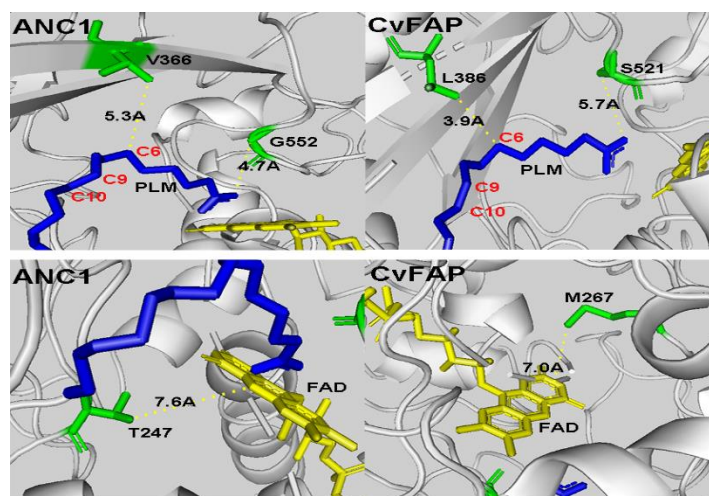


Figure 38. Structure of active site of sANC1 and CvFAP with palmitic acid. PLM shown in blue, and FAD is shown in yellow. The yellow dashed showing the different minimum distances between PLM / FAD and “hot” positions of sANC1 and CvFAP. In sANC1 was V366, G552, and T247, as shown in green; in CvFAP was L386, S521, and S267, as shown in green. Structures were visualized with PyMOL.^[123]

Based on the prediction, the three variants of sANC1 were generated by site-directed mutagenesis *via* QuikChange protocol, and the desired point mutations were introduced in the sANC1 gene using pET28a-sANC1 as a template. After confirming the success of obtaining desired mutations by Sanger sequencing, all three mutation-constructed plasmids were transformed into *E.coli* BL21(DE3). Three variants were successfully soluble expressed (Figure 39).

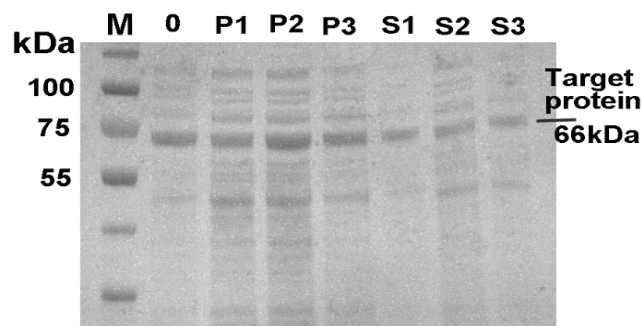


Figure 39. SDS-PAGE analysis of sANC1 and variants (MW=66 kDa) expression. All variants could be soluble express. PageRuler Prestained Protein Ladder (Thermo Scientific) was used as standard.

With palmitic acid as substrate, sANC1_{V366L} specific activity was 3-fold greater than for sANC1_{WT} (3.7 and 1.2 U/mg), sANC1_{T247M} activity (1.3 U/mg) was similar as sANC1_{WT}, sANC1_{G552S} activity (0.52 U/mg) decreased 2-fold comparing with sANC1_{WT}. From these results, we identified that residue V366 was a potential contributor to improve sANC1 activity. The leucine side chain is larger than valine, and the substrate might be better stabilized through hydrophobic interactions. The introduced G552S in sANC1 caused the decrease of specific activity, and this indicates that comparing the residues nearby the FAD binding site between CvFAP and ANC1 might reveal the photodecarboxylation mechanism in FAP.

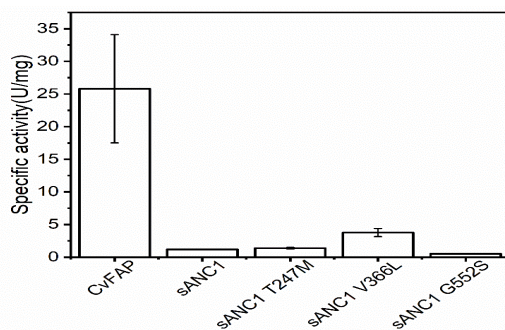


Figure 40. The specific activity of sANC1 mutants

3.11. Rapid Nile red quantification method for pentadecane production

In view that CvFAP catalyzes fatty acid substrates to generate alkenes, either fatty acid consumption or alkene production could be selected as the means to evaluate enzyme activity. The only readily detectable approach for fatty acid and alkanes was the highly specified and sophisticated GC approach, which needs complicated preparative steps, i.e., extraction, purification, and concentration. Consequently, the engineering study of catalytic activity or stability of FAPs can be extremely energy and time-consuming. A faster and more convenient method to accelerate the quantification process would be ideal.

Fluorescence spectroscopy is a simple, rapid, and sensitive detectable technique used as combinatorial and high throughput experimentation techniques. As shown in Figure 41, Nile red is a benzo phenoxazine dye with the properties of a fluorescent hydrophobic probe. The dye fluoresces in the presence of hydrophobic compounds (*i.e.*, fatty acids, membrane lipids),^[94] and the fluorescence character shifts depending on the hydrophobicity of the compounds and their surroundings.^[95] Such as Nile red with neutral lipids, the maximum emission wavelength is shorter, and the fluorescence intensity also higher than with polar lipids.^[95] Lipids composed of unsaturated fatty acids presented more vigorous fluorescence intensity than saturated fatty acids.^[96] *Spencer et al.* presented a high throughput screening for intracellular lipid droplets based on the hydrophobic fluorescence dye Nile Red.^[97] Moreover, the fluorescence intensity of Nile red scales with its hydrophobic environment,^[98] it was used to evaluate lipid content of bacteria,^[99] microalgae,^[100] and yeast.^[94]

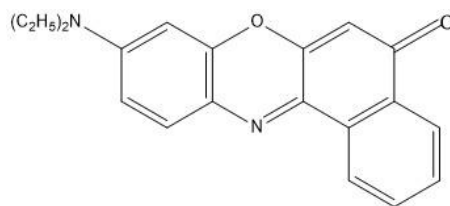


Figure 41. Chemical structure of Nile red. The dye is hydrophobic and fluorescent.

In our case, FAPs enable catalysis decarboxylation of fatty acids to produce alkanes. We reasoned the Nile red with fatty acid and alkane will present different fluorescence characters. The fluorescent intensity could be used to evaluate the concentration of fatty acids and alkenes. Thus, as the flowchart shown in Figure 42, we aim to develop a fluorescent assay applied to rapid screening processes and quantification of alkene production: 1. Create the variants (*e.g.*, directed-mutagenesis, site-saturation mutagenesis) of desired properties. 2. Inoculate the mutant strains in 96-deep-well plates, and induce the cells individually. 3. Perform decarboxylation reaction. 4. Background content assessment. 5. Adding Nile red and measuring the emission spectrum, screening and identifying improved alkene producing mutants.

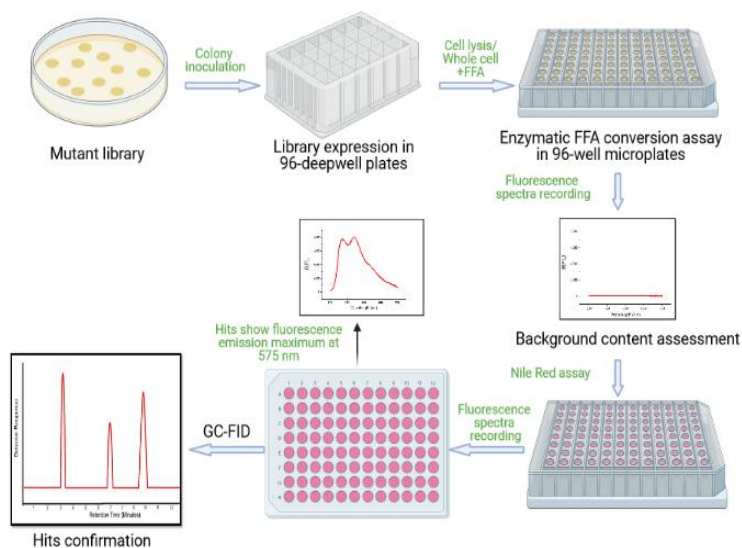


Figure 42. Flowchart of the 96-well plate format high-throughput screening strategy via Nile red.

We found that the selectivity for pentadecane staining is obtained when excitation wavelength at 450 nm. Pentadecane presented distinguish spectrum peaks at 570 nm when the emission spectra from 500 nm to 700 nm (Figure 43). Palmitic acid showed a low spectrum peak around 650 nm, the Nile red and the Nile red in Tris-HCl (pH 8.5, 100 mM) Buffer showed no spectrum peak. We choose to increase the concentration of palmitic acid from 1 mM to 3 mM or even 6 mM to observe whether the fluorescence intensity would be varied. However, fluorescence intensity did not change (Figure 43).

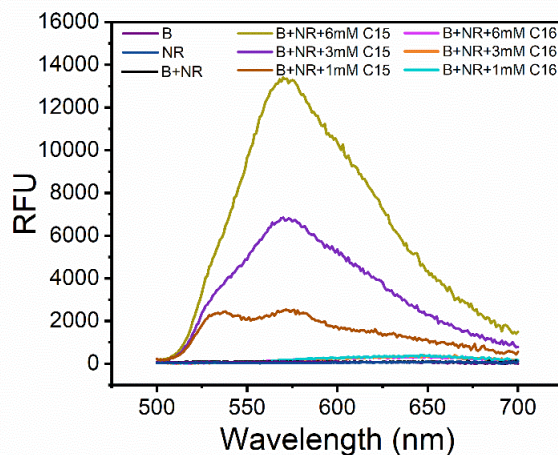


Figure 43. Emission spectra of Nile red-stained pentadecane and palmitic acid. Emission spectra were recorded with an excitation wavelength of 450 nm. *B: Tris-HCl (pH 8.5, 100 mM) Buffer + 30% DMSO; NR: Nile Red; C15: Pentadecane; C16: Palmitic acid.

The fluorescence of pentadecane staining with Nile red showed an emission spectra peak at around 570 nm. Further optimization to reduce scattering signal attempted via decreasing DMSO content. The scattering signal at 530 nm decreased when the DMSO volume ratio in the reaction system was reduced from 30% to 10% (Figure 44). We assumed the scattering signal at 630 nm was from ethyl acetate due to pentadecane only could be solved in ethyl acetate. Nevertheless, ethyl acetate will not be an influencer when reaction only performs at reaction buffer containing co-solvent DMSO. The assays were run twice, with each reaction set to determine the reproducibility of the procedure. The results showed that the assay is highly reproducible.

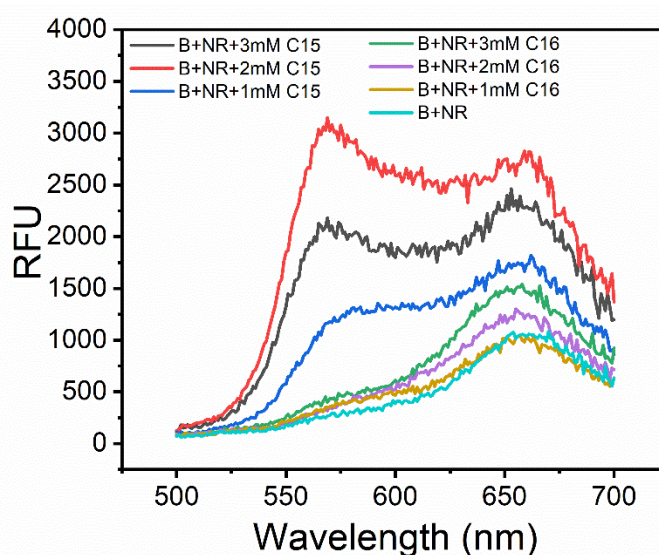


Figure 44. Emission spectra of Nile red-stained pentadecane and palmitic acid.

Emission spectra were recorded with an excitation wavelength of 450 nm. B: Tris-HCl (pH 8.5, 100 mM) Buffer + 10% DMSO; NR: Nile Red; C15: Pentadecane; C16: Palmitic acid.

Therefore, further study was carried out to determine whether Nile red is a suitable reagent for identifying and determining pentadecane when cells or cell-free extract are added in the reaction system.

A brief treatment of cultured cells and the cell-free extract was described in methods. Comparing with the blank buffer system, the cell or cell-free extract exhibits a natural fluorescent peak at 530 nm without Nile red staining (Figure 45). After staining with Nile red, the lipid of membranes or other hydrophobic domains of cellular protein showed a fluorescent peak at 630 nm, but in all cases. At 570 nm, we observed no intensive fluorescence signal from stained cells and cell-free extract. When added 1 mM pentadecane in the systems containing cells or CFE, the emission maxima peaks of pentadecane at 570 nm was observed, and the intensity directly corresponds to the concentration of pentadecane.

Consider the unique fluorescence properties of CvFAP and the formation of hydrophobic domains of cellular protein that may influence the fluorescent intensity. The ancestral reconstruction enzyme ANC1 showed a higher soluble expression level and strong fluorescence intensity. The experiments presented whether the Nile red staining method suitable for ANC1 was presented in Figure 45. Apparently, under the conditions employed as CvFAP, the Nile red can be a general stain for analyzing pentadecane production in photodecarboxylation reaction.

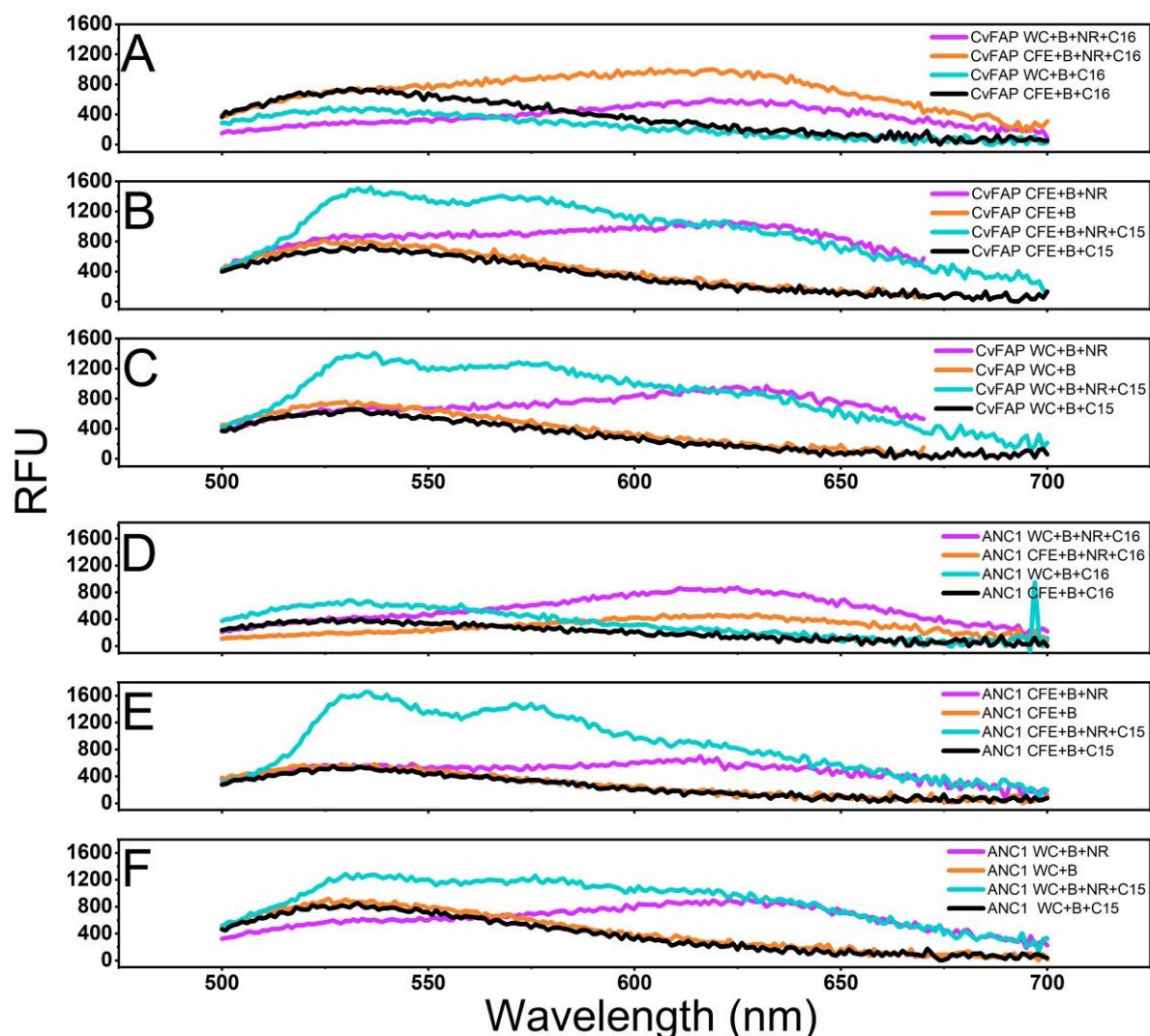


Figure 45. Fluorescence emission spectra of Nile red-stained cell-free extract and whole-cell of FAPs.

Emission spectra were recorded with an excitation wavelength of 450 nm. B: Tris-HCl (pH 8.5, 100 mM) Buffer + 10% DMSO; NR: Nile Red; C15: Pentadecane; C16: Palmitic acid; CFE: cell free extract=0.5 mg mL⁻¹; WC: whole cell=0.5 mg mL⁻¹.

This part of our study demonstrated the utility of Nile red as a stain to detect pentadecane by fluorescence-based assay would be the promising high throughput screening method

for selecting the target in mutations library of FAPs. The assay could be used in whole-cell and cell free extract systems. We have not optimized our method for efficiency, but the procedure is significantly faster than analysis by GC.

4. Conclusion and outlook

With 3.5 billion years of evolution, thousands of different enzymatic activities have evolved in living cells. Most enzymatic reactions are thermally activated, whereas light-driven reactions have seldom been found in nature. For an extended period, only two families of natural photoenzymes were found, the photolyases involved in the repair of DNA damage in many organisms^{[48]–[51]} and the light-driven protochlorophyllide reductases involved in chlorophyll synthesis.^[52] The chlorophyll f synthase was considered the other type of photoenzymes, but the light-force nature of the reaction was not established.^[53] Therefore, photodecarboxylases FAP was considered as the third novel type of photoenzymes.

In our study, we demonstrated that the expression and catalytic activity of CvFAP were influenced by the used plasmid backbone and inducer concentration. With pET28(a) plasmid, when protein was induced with 0.5 mM IPTG, the expression level and catalytic activity were improved compared with 0.1 mM IPTG. However, CvFAP produced with 0.5 mM IPTG seems to be approximately three times less thermostable than CvFAP produced with 0.1 mM IPTG. When pBAD plasmid was used, the expression level was 5-fold improved, but the catalytic activity was similar comparing with CvFAP produced with 0.1 mM IPTG.

The most exciting achievement in this work was the reconstruction of possible ancestral decarboxylases from a set of 12 extant sequences that were classified under the fatty acid decarboxylases clade within the glucose-methanol choline (GMC) oxidoreductase family. One of the resurrected enzymes (ANC1) showed activity in the decarboxylation of fatty acids.

We observed that the denaturation of CvFAP and ANC1 yielded pronounced decreasing fluorescent intensity, which was different from other flavoproteins. As a consequence of the promising prognoses, during the denaturation of flavoproteins, the secondary and tertiary structure of the protein is disrupted, and interactions with flavin break down, usually leading to significant increases in the flavin fluorescence. From the observation of CvFAP, a blue-shift in the FAD peak fluorescent feature and the decrease in FAD fluorescence emission signal after denaturation was observed. For ANC1, no fluorescent shift was observed. Instead, the fluorescent intensity was much stronger than normal flavoprotein. This indicates that as a photoenzyme, the FAD cofactor's environment/electronic fluorescent properties in FAP were unique. The residues nearby the FAD binding site might be essential to functional charge transfer. Thus, further studies about the difference of residues between CvFAP and ANC1 might contribute to the mechanism study of FAP.

As the recently discovered photoenzymes, the understanding of structure-function relationships of FAPs was deficient. The only approach for uncovering and improving properties of FAPs was structure-based engineering of CvFAP. In our case, the sequence identity between CvFAP and sANC1 is 76%, but this robust ancestral enzyme demonstrated a 12-fold enhanced purification yield, as well as better stability and folding efficiency compared with the extant CvFAP form. We observed that most of the

substitutions on the outer surface of ANC1 were towards more hydrophilic amino acids, which improved the soluble expression of ANC1. There are more rigid regions in ANC1 comparing with CvFAP. The rigid regions in ANC1 might tightly interact with other amino acids, which explains that ANC1 is more stable than CvFAP. Thus, ancestral photoenzymes provided another example for understanding the mechanism of photodecarboxylase and engineering templates for obtaining FAPs with desired properties. The different selectivity was verified when the predicted chloroplast sequence in ANC1 was removed. The variant sANC1_{V366L} presented 3-fold higher specific activity than sANC1_{WT}.

Thereby, the limitation of synthetic applications of CvFAP attributed to its instability and low protein production was overcome by reconstructing ancestral photoenzyme. Furthermore, introducing deep eutectic solvents (DES) as an alternative solvent system provided another approach to broaden the application of FAPs. Our results demonstrate that the DESs system enables maintenance and improvement of photodecarboxylation activity of FAPs. On the other side, the fatty acids used as substrates could be utilized as hydrogen donors for generating DESs. Thus, the issues (e.g., poor solubility of fatty acids and limitation of loading substrates) could be solved. Such the C8: DL-menthol mixtures system produced 1.3 mM Heptane when there was no production of heptane in the aqueous system.

Another achievement in this study was generating a fluorescence-based assay for detecting and quantifying pentadecane production. The fluorescent dye Nile red with palmitic acid and pentadecane presented different fluorescence characters. The fluorescent intensity could be used to evaluate the concentration of pentadecane. The assay could be used in whole cell and cell free extract systems. We have not optimized our method for efficiency, but the procedure is significantly faster than analysis by GC. Using this assay, we would screen mutations with better capacity in a large number of mutations for a short period.

5. Experimental section

5.1. Sequences, plasmid, strains, and growth conditions

All DNA sequences used in this study are collected in Table 12. All the ancestral sequence reconstructions were synthesized and cloned into pET28a(+).

Table 12. DNA sequences were used in this study.

DNA sequence (5' to 3')
<p>CvFAP ATGCAGCTACGGTCGCCGGTGCAGCAGCAGATGCACCAATGGTTGCTTTCCC CGTCAGCAGAGCTGCTGCGCGTTTCAGCGATCATAACTACCGGCGCACCGGTC TGACCACCCGGAATTTTCGGAACAACGCTAGCGTCAACAACGCGCAGGCCTT CAACACCGTGAACACGCAGCTGGTTGTCTACCACAGAGCTGCTGTCACCTGC GTTACCCATTTTACAGGTGCCAGTGATAGCGTTGGACGAGTGGATAGAACGAC GGATATATTCATCGATCTGATCATCAGAAACAACGCCGCTACCTGGGAACAGC TCACCATCCAGGTATTCGGACAGAGCAGAGCTACGCGCAACATCACGTGCC AATGGATGCCTTTACGCAGGGTAGCCAGATCAGCACCGTCTTTGTCCGGTCAG GTAACCTGGTGACAGTTTCGGCGGCGCAAACGGGTCAGCGGATTTAAGACCG ACGGAGCCGGTAGACTGCGGACGGCAAGCGATCAGCTGCATGGTGATGCCG CTCGGCCATTTTCAGACCCTGGCTCTGGAATTTAGCAAACGAACGTAGGTGCT AACACCGTCCGGGTCCAGCGCCATACCTGGAACGAAGCGAACCTGCAGGTCC GGCAGCGCCTGACCCGCGGTACGAACGAAGGCACCGCGATCGCAACCGGTG GAAGTCAGACCGCCACGACCACCCAGCAGGTAGGATGCGATTGCACGTTTAC GGATCTGGCCTTTTTTCGTTGTAGATGTGATCAGAAATGGCAATACCGTCGTAT TTTTCTTTAACCGGAGCCGCGGTCAGGCACGCCGGCTGATCCTGCAGGTTCT GGCCAACACCAGCCAGGTTGCTAACAACCGGGATGCCGAATTTCTTCAGCTC AGCAGACGGGCCAACGCCGAATGTTTCAGCAGGAACGGGGTGTGAACAGC ACCTGCGCACATGATGACCTCACCACCCGGAGCCAGTTCCGCAGACAGGCGT TCGCCGGTTGGGCCGTCGGTGGAGAATTCAACACCCAGAGCCTGCGCTTTGC CCGCAGCCTGGTTCGATGTTGACTTTGGTCACTGCAGCGCCGGTACGTACCTG CAGGTTGCGACGACCCAGCACAGGTTTCAGATACTGACGGTACATATCCGCG CGGGTGCCTTTATCCTGCATCACCTGAAAGGTGCCGTAACCGGCGTGGTCAT GGCTCCAATCGTTGAAATCGGAGTTCGGGGTAAGACCAACTTCTTCAGCAGC CTTGAAGAAAGCAGTGTGCAGCTGTTTGTGGTGTAACGCGGGTTTTCCACAC GCATCGGGCCGCGCTGCCATGATAAGCGCCCGGACCGAAGTCCGCGTTGG TTTCCGCTGGACGAACCAAGACAGAACGTCTTCGCTGGACCAGCCTTCAAC ACCCCATGCGTCGTAATCACCCGCCGCACCACGGTGGTACAGAGTGGCGTTA GTCGCGCTGGAACCGCCAGCAGACGGCCACGCGCCATGTAGATCTGACGT TCCGCAAGCTGTTCTGCAAGTTCAGAGAACAGGTTCCAGTCCAGCGGGGAGC GGAACAGGCGGGTGATCGCCGCCGAATCTTAACGTCGCGGGAGGTGTTAT CCGGGCCTGCTTCCAGAACCAGTACACGTTTGGAACCGTCAGCGCTCAGACG GTTTGCCAGCACGCACGCCGCGGTGCCACCGCCAACCAGGATGTAGTCATAT TTCTGACCCGCCACCGGAGACGAAGAATCGGACAGGACTTTACGGATGTCTT CAACGGCAGACGCGGATTGGAAGTACAGGTTTTCTCGATCCCGGCCAGGTT AGCGTCGAGGAACTCTTCAACTGACCTTTAGACAGTGCACCCACTTTGGTTG</p>

CCGCCACTTCACCGTTTTTTGAACAGCAGCAGAGTCGGGATACCACGGATGCC
 ATATTTTCGGCGCAGTGCCAGGGTTTTGATCGATGTTTCAGTTTTGCAACGGTCA
 GTTTGCCCTGATATTCGTCAGCGATTTTCATCCAGAATCGGGGCGATCATTTTG
 CACGGACCGCACCACTCTGCCAGAAATCGACGAGGATCGCCCCGTCCGCTT
 TGAGTACATCCGTGTCAAACCTGTCGTCAGTCAGGTGAATAATTTTATCGCTCA
 TAGAAGAACCATGGTGTGATGGTGTGAGAAGATTTTCAT

ANC1

ATGATGTCCTCTCGTTCGGTACTGCTGGGTGCTCGTCCAGTTACTCGTGCACC
 ACTGCCAGCTGCAGCTCCTACTCCTGCACGTCGTTCCCAGTCCGGTGCTGCA
 CGTCGTGCTCTGTCTGCTCGTGGTGGTCTGGGCGAAACCCTGCGTAGCGTTG
 TTAGCGCGTCTAGCTCCCCGGTGGCTTCCGAAAATACGATTACATTCTGGTA
 GGCGGGCGGTACTGCAGGCTGTGTACTGGCAAATCGTCTGAGCGCTGACGGTT
 CCAAACGTGTACTGGTTCTGGAAGCTGGCCCAGCGAACAAGTCCCGTGAAGT
 GCGTATTCTGCAGGTATTACTCGCCTGTTCCGCAGCCCCGCTGGATTGGAATC
 TGTATAGCGAACTGCAACAGCAGCTGGCCGAACGCGAAATCTATCTGGCTCG
 TGGTCGTCTGCTGGGCGGTTCTTCTGCCACTAACGCTACCCTGTATCACCGTG
 GCACTGCTGCAGACTATGACGCGTGGGGTGTAGAAGGCTGGACCTCCGAAGA
 CGTGCTGTCTTGGTTCGTTAAAGCAGAAAACAATTACGGCGGTGGTCCGGGT
 GCATACCATGGTACTGGCGGTCCGATGCGCGTAGAAAATCCGCGCTATCAGA
 ACCCACTGCATGAGGCTTTCTTCAAGGCAGCGGAAGAAGCAGGTCTGCCGGC
 TAACCCGGATTTCAACGACTGGTCTCACCCCTCAGGCTGGCTACGGTGAGTTC
 CAGGTGACCCAACGTAAAGGCCAGCGTGCTGACACCTACCGTACCTACCTGA
 AACCGGTGATGGGTGCTTCTAATCTGCAAGTTCTGACCGGTGCTGCGGTTAC
 CAAAGTTAACATCGAAAATCTGGCGGTGGTCCCTCGTGCACGTGGTGTGAAT
 TCCAGGCTAACGGTCAGGATGGTGACCGTCACTCTGCCGAACTGGCGCCAGG
 TGGTGAAGTGCTGATGTGTGCCGGTGCAGTACACACCCCGCATCTGCTGATG
 CTGTCTGGTATCGGTCCAGCAGCTGAGCTGCGTGAACACGGCATCCCGGTTG
 TTAGCGACCTGCCGGGTGTTGGCCAGAACCTGCAGGACCATCCGGCTTGTGT
 CACCGCTGCGCGTGTTAAGGAAAAGTATGAACCGATCTCTGTTACCGACGAG
 ATTTATAACGAAAAAGGCAAAATCCGTCCCTCGTGCTGTAGCACAATACCTGCT
 GGGCGGCCGTGGTCCCTCTGACTTCTACTGGTTGCGATCATGGTGCCTTTGTA
 CGCACGGCAGGCCAGGCACAACCTGATCTGCAGATCCGTTTTGTACCTGGTC
 TGGCTCTGGACCCGGATGGTGTCTTCTTATATCGCGTTCGGCAAGCTGAAA
 AGCCAGGGTCAGAAATGGCCGTCTGGTATTACTCTGCAGCTGCTGGCTGTGC
 GTCCGAAATCCAAAGGCTCCGTTGGCCTGAAAAGCGCTGATCCGTTTGCGCC
 ACCGAAAATCGACATTGGTTATCTGACTGACAAGGCTGATCTGGCAACCCTGC
 GTAACGGCATCAAACCTGGCGCGCGAAATTGCTGCCAGCCTGCACTGGGTGA
 ATACCTGGGCGAAGAAGTGTCCCGGGGCGCAGCAGTATCTTCTGACGAGGAA
 ATCGACGAATATATTCGCCGTAAGTGTGCACAGCGGCAACGCTCTGGTCGGTA
 CCTGCCGTATGGGTAAACGCGTCCGATAGCTCTGCGGTGGTTGACAGCGAACT
 GCGCGTCTTCGGCGTTGAAGGTCTGCGTGTGCGTGGATGCAAGCGTTATTCCA
 CGCATTCCAGGTGGCCAGACCGGTGCTCCTACCGTAATGGTTGCGGAACGTG
 CTGCCGCACTGCTGCGT

ANC2

ATGATGTCTGCACAGTCCGTA CTGCTGGGTGCGCGTCCTGTTACCCGTGCTC
 CTGTACCAGCAGCAGCACCGACTCCAGCACGTGCTAGCCAGAGCTCTGCTGC
 TCGTCGTGCTCTGTCTGCACGTGGTGGCGTAGGTGAAACGCTGCGTTCTGTC
 GTTCGTGCATCCTCTTCCCCGGCAGCTTCTGAAAAATACGATTACATCCTGGT
 TGGCGGTGGCACCGCTGGTTGTGTTCTGGCAAACCGCCTGAGCGCAGATGGT
 TCCAAGCGTGTTCTGGTTCTGGAAGCTGGCCCCGGCTAACAAATCCC GCGAAG
 TTCGTGTTCCGGCGGGTATTACCCGTCTGTTTAAATCTCCGCTGGATTGGAAC
 CTGTA CTCCGAACTGCAGCAGCAGCTGGCAGACCGTGAAATCTATCTGGCCC
 GTGGTCTGCTGCTGGGTGGTTCTTCTGCGACCAACGCTACGCTGTACCACCG
 TGGTACTGCCGCTGATTATGACGCGTGGGGTGTGGAAGGCTGGTCTAGCGAA
 GATGTTCTGTCTGGTTCGTGAAAGCTGAAAACAATGCAGCCGGTGGTCCTG
 GTGCTTACCATGGTACCGGTGGTCCAATGCGTGTGAGAACC CGCGCTACCA
 GAACCCGCTGCATGAAGCGTTTTTCAAAGCTGCGGAAGAAGCAGGTCTGCCG
 GCAAACCCAGATTTCAACGACTGGAGCCACCCGCAGGCTGGCTATGGTGAAT
 TCCAGGTCACCCAGAAAAAGGGCCAGCGTGCTGACGCGTACCGTACGTACCT
 GAAACCGGCCATGGGTCTGAGCAACCTGCAGGTTCTGACTGGCGCTGCGGT
 GACCAAAGTGAACATTGAAAAGAGCGGTGGTGGTGC GCGTGCACGTGGTGTGTA
 GAATTCCAAGCTAACGGCCAGAACGGTGAACGTCACTCTGCCGAACTGGCAC
 CGGGTGGTGAAGTTCTGATGTGTGCCGGTGCGGTGCATACTCCACACCTGCT
 GATGCTGTCTGGCATCGGTCCAGCGGCTGAGCTGCGTGAACACGGTATCCCG
 GTTGTTAGCGACCTGCCGGGTGTTGGTCAGAACCTGCAGGATCACCCGGCAT
 GCGTTACCGCGGCTCGTGTTAAAGAGAAATATGAACCTATCAGCGTTACCGAC
 GAAATTTACAATGAGAAGGGTAAAATTCGTGCACGTGCTATCGCTCAATACCT
 GCTGGGCGGTCTGGTCTCTGGCTTCTACTGGTTGTGACCACGGCGCATT
 GTGCGTACCTCCGGCCAGTCTCAACCAGATCTGCAGATTCGTTTTGTCCCGG
 GTCTGGCTCTGGATCCAGATGGTGTCTCTTCTTATGTGGCCTTCGGTAAAATG
 AAAAGCCAGGGTTCGTAATGGCCGTCCGGCATCACGCTGCAGCTGATCGCTT
 GTCGCCCTAAGTCCAAGGGTTCTGTTGGTCTGAAATCTGCCGATCCGTTTGCT
 CCGCCAAAATCGATATCGGTTACCTGACTGATAAGGCTGATCTGGCTACCCT
 GCGTAACGGTATCCGCCTGGCACGTGAAATCGCTGCTCAGCCAGCGCTGAGC
 GAATACCTGGGTGAAGAACTGTTCCCGGGTGCGGCAGCGTCCTCTGATGAAG
 AAATCGATGAATATATTCGCCGTACCGTGACAGCGGTAACGCACTGGTGGG
 TACTTGCAAAATGGGTAACGCATCCGATAGCAACGCCGTTGTAGATTCTGAAC
 TGCGTGTGTTCCGGCGTTGAGGGTCTGCGCGTAGTAGACGCGAGCGTAATCCC
 GAAAATCCCTGGTGGTCAGACGGGTGCACCGACCGTTATGGTAGCTGAACGT
 GCGGCGGCTCTGCTGCGT

ANC3

ATGATGTCTCGCATTGCAGTACTGCTGTCTACCAGCCCATCCCTGATTGCAGC
 TTCTTCTAGCAGCTCTCCAGCTCCGGCACGTGCTCCAGCATCTAGCTCTGCGC
 GTCGTGCTCTGTCTCTGCGTGGTGGTGCCGGTACTGTTGCGGCAAGCGTTGCG
 TCGTGCTGCACCGTCTCCGGTAGCGTCTACTACCTATGATTACATTATCGTCG
 GTGGCGGCACTGCAGGCTGTGTTCTGGCAAACCGTCTGTCCGCGGATGGCTC
 TAAACGTGTTCTGGTTCTGGAGGCTGGCGGTTCTGCCTCTCGCTCTCTGTATG
 TTCGTATCCCGGCTGGTATCACCCGTCTGTTTCGTTCTGCATATGACTGGAAT
 TTCAGCACCGAACC GGAGCAGGGCCTGAATGGTTCGTGAAGTTTACCTGTGCC
 GCGGTAAAGCTCTGGGCGGTAGCTCTTGTACTAACGTCATGCTGTACCATCGT

GGTTCTGCCGCAGACTACGACGCTTGGGCTGCTGCTGGTGCTGAAGGCTGG
GGTCCAGAGGATGTTCTGCCGTATTTAAGAAAGCAGAAAACAACGCTGCGG
GTGGCGCAAATCAGTACCATGTTCCGGTGGTCCGCTGGCAGTTGAAGATGT
ACGCTACCAGAACCCGCTGTCCAAGGCGTTTCTGGAAGCAGCCGAAGAAGCA
GGTCTGCGTGCAAACCCGGACTTCAATGACTGGTCTCACCCACAAGACGGTT
ACGGCCGTTTCCAGGTAACCTCAGCGTAACGGCCGTCGTTGTTCTGCCGCAAC
TGCATATCTGCGTGCAGCTCGCGGTTCGTTCTAACCTGCATGTTGTTACCGGCG
CTGCTGCCACCCGTGTAACCTCTGGAAGGTTCTGGTGATGGTGCTGGTGGTGG
TGGTGGCGGTAAACTCGTCCATGGACTGGTCCGGCCGTAACCGGTCAAGCT
GGTCCTCGTGCTACGGGCGTTGAATTCATTGGTGCAGATGGTGACCGTCGTA
CCGCCGAAGTCTCCTCTGCACGTCTGGCACAGGGTGGCGAGGTTCTGCTGTG
TGCGGGTGCAGTTCATTCCCCGCAGCTGCTGATGCTGTCCGGCATTGGCCCG
GCTGCACATCTGCGTGAAGTAGGTATCCCGGTTGTGGCTGATCTGCCGGGTG
TTGGTCAGAACCTGCAGGATCATCCGGCAGTCGTGGTAAGCTATCGCAGCAA
AAAAGGTGTGAGCGTGACCGATGAAATTCGCCTGTTCCGGTACTTCTAAAACCA
ACCCGATGGCAGTTCTGCAGTGGCTGCTGTTCCGGTCGTGGCCCACTGACTTC
TCCTGGCTGCGATCACGGCGGCTTTGTGCGTACCTCTCCATCTCTGGAGCAG
CCGGATCTGCAGATCCGTTTCGTACCTGCTCGTGCTCTGGATCCGGATGGTAT
GTCTACTTACACGGAGTTCGGCACTGCGGCGAAGCGTCTGTCTGGTTTCACT
CTGCAGTCTGTTGCGTGTGTCGTCGAAAAGCAAAGGTCGTGTTTCGTCTGGCTTC
TGCGGATCCGTTCCGCAAACCGATGATTGAAGGCGGTTACCTGTCTGATGAA
GCAGATCTGGCGACCCTGCGTAATGGTATTCGTCTGGGTTCGTGAGCTGGCTG
CCCAACCAGCATTGCGCAATACCGTGGCGAGGAAGTATTCCCGGGTGCTGC
AGTTCAGTCCGATGAGGACATCGACGCCTATATCCGTAACACCGTACACACCG
CTAACGCTCTGGTAGGTACCTGTCGTATGGGTAATGCGAGCGACCGTTCCGC
AGTTGTAGACCCAGAACTGCGCGTAATTGGTGTGGCGGTCTGCGCGTTGTA
GACGCGTCCGTAATGCCTACCATCCCGGGTGGTCAGACTGGTGCTCCGACCG
TAATGCTGGCTGAACGTGCTGCGGATCTGGTCCGT

FVPGMALDPDGVSTYVRFQAKFQSQGLKWPSGITMQLIACRPQSTGSVGLKSADPFA
 PPKLSPGYLTDKDGADLATLRKGIHWARDVARSSALSEYLDGELFPGSGVVSDDQID
 EYIRRSIHSSNAITGTCKMGNAGDSSSVVDNQLRVHGVEGLRVVDASVVPKIPGGQT
 GAPVVMIAERAAALLTGKATIGASAAAPATVAA

Alcohol oxidase [*Coccomyxa subellipsoidea* C-169]

MMASQSVFLGTRPATRSPLPIGRAGHGSAGRRALRVRAIIKSDNPAADKYDFILVGG
 GTAGCVLANRLTADGSKKVLLEAGGANKAREVTRTPAGLPRLFKSALDWNLYSSLQ
 QAASDRSIYLARGKLLGGSSATNATLYHRGTAADYDAWGVPGWTSQDALRWFQAE
 NNCRGIEDGVHGTGGLMRVENPRYNNPLHEVFFQAAKQAGLPENDNFNNWGRSQ
 AGYGEFQVTHSKGERADCFRMYLEPVMGRSNLTVLTGAKTLKIETEKSSGATVSRG
 VTFQVNGQDGSKHSAELAAGGEVVLCAAGSIHSPQILQLSGIGPQAE LRSKDIPVADL
 PGVGNMMDHPACL SAFYLKESAGPISVTDELLHTNGRIRARAILKYLLFKKGPLATT
 GCDHGAFVKTAGQSEPD LQIRFVPGALDPDGIGSYTAFGKMKDQKWPSGITFQLLQ
 VRPKSRGSVGLRSDDPWDA PKLDIGFLTDKEGADLATLRSGIKLSREIAAEPAFGAYV
 GNELHPGAAASSDSAIDSFIRD TVHSGNANVGTCSMGVNGNAVVDPSLRVFGIRGLR
 VADASVIPVIPGGQTGAATVMVAERAAEILLGSNQKQPAAVPAAQPALA

Predicted protein [*Chlamydomonas reinhardtii*]

MMLGPKTVTRGATKGAAPRSMAARRVGGARRLSVRAAAGPAGSEKFDYVLVGGGT
 ASCVLANKLSADGNKKVLVLEAGPTGDAMEVAVPAGITRIFAHPVMDWGMSSLTQK
 QLVAREIYLARGRMLGGSSGSNATLYHRGSAADYDAWGLEGWSSKDVLDWVKAEC
 CYADGPKPYHGTGGSMNTEQPRYENVLHDEFFKAAAATGLPANPDFNDWSHPQDG
 FGEFQVSQKKGQRADTYR TYLKPAMARGNLKVVIGARATKVNIEKGSSGARTTGEV
 YAMQQFGDRFTAELAPGGEVLMCSGAVHTPHLLMLSGVGPAAATLKEHGIDVVS DLS
 GVGQNLQDHPAAVLAARAKPEFEKLSVTSEVYDDKCNILGAVAQYLFQRRGPLATT
 GCDHGAFVRTSSSLSQPDLQMRFPVPGCALDPDGVKSYIVFGELKKQGRWPGGITL
 QLLAIRAKSKGSIGLKAADPFINPAININYFSDPADLATLVNAV KMARKIAAQEPLKKYL
 QEETFPGERASSDKDLEEYIRRTVHSGNALVGTAA MGASPAAGAVVSSADLKVFGV
 EGLRVVDASVLPRI PGGQTGAATVMVAERAAALLRGQATIAPSRQPVAV

Hypothetical protein GPECTOR_35g846 [*Gonium pectorale*]

MMLGRKPVAPAKGASAARTVRPVRLAGGRRQLVVSAAAAPVDPAEKYDYILVGGGT
 AGCVLANKLSADGNKKVLVLEAGPSGDSLEVAVPAGIARLFAHPVMDWGMSSLTQK
 QLVAREIYLARGRLLGGSSGTNATLYHRGTSSDYDSWGLEGWTSKDVLDWVKAEC
 YGDGPKPYHGNSGSMNVEQPRYQNPLHEEFFRAAAAAGIPANPDFNDWSRPQDGY
 GEFQVAQNKGQRADTYR TYLKPALS RGNLKVVTGARTTKVHIEKGSSGPRARGVEF
 ATQQFGDRYSAQLAPGGEVLMCTGAVHTPHLLMLSGVGPAAALREHGVDVADLA
 GVGANLQDHPAAVVAVRAKPEFEKLSVTSEIYDEKCNILGAVAQYLFNRRGPLATT
 GCDHGAFVRTSGSHSQPDLQMRFPVPGCALDPDGVKSYIVFGELKKQGRWPGGITL
 QLLAIRAKSKGSIGLKAADPFINPAININYFSDPADLATLKQGV RMARDIARQEPLRKYL
 QEETFPGERASSDSIEEYVRRTVHSGNALVGT CAMGTSPAKGAVVSSSDLKVFGV
 EGLRVVDASVLPQIPGGQTGAATVMVAERAAALLKGQTTMAPSRQPVA

Hypothetical protein VOLCADRAFT_103673 [*Volvox carteri f. nagariensis*]

MLLGQRPF GAPAKGAMPCWKAARHGGVAGVARRPVAVKAAASVGSEKFDYILVGG
 GTAGCVLANKLSANGSKKVLVLEAGPTGDAMEVAVPAGIARLFAHPVFDWGMSSLT

QQQLVAREIYLARGRLLGGSSGTNATLYHRGTPADYDSWGLEGWTSKDLLDWFVK
 AECYGDGPRAFHGQSGSMNVEQPRYQNVLHDEFFRAAAAAGLPANEDFNDWSRP
 QEGYGEFQVAQKNGERADTYRXYLKPAMGRDNLKVMGTGARTTKVHIEKSSTGPRA
 RGVEYATQQFGERYTAELTPGGEVLMCTGAVHTPHLLMLSGIGPAPTLEHGLDVIS
 SLPGVGANLQDHPAAVLAVRAKPEFEGLSVTSEIYDSKCNIRLGAVMKYLFGRGPL
 ATTGCDHGAFVRTSASHSQPDLQMRFPVPGCALDPDGVKSYIVFELKKQGRAWPG
 GITLQLLGIKAKSRGSLKAADPFINPAININYFSDPEDLATLKNVRIAREIVAQEPLR
 KYLLEETFPGERANTDKDIEEYVRRTVHSGNALVGTGAMGTPASGAVVSSADLKVF
 GVDGLRVVDASVLPRIPIGGQTGAATVMVAERAAAMLLGQATITSRREPAAV

Hypothetical protein AURANDRAFT_59047 [*Aureococcus anophagefferens*]

MGRTLVLKVATTSYDYIIAGGGTAGCVLANRLESDPSKKVLVLEAGDRGPN SPLVKIP
 VAILKLFKSAYDWNFATRPSEAVADRSLYVCRGKGLGGSSLTNVMLYNRGSANDYD
 AWAAACGDDSWGAEEMLYFKKAEDCLVPAHRANHYHGVGGPYASSHVPYTNEM
 STAFVEAAVEDGGVRNGDFNDWSTSQVGFGRFAVSQRKGARVDAATAYLPRKVRR
 RANLDVVRGAALSGVTWNANKATGVEFAFGGVSGIACGGEVILSGGAVHSPQMLML
 SGVGAQAQLEEFQIPVADRPGVGNLQDHPACLVSWRGSAKAQKSHSTQLRIPG
 TTKTSPKALLQWLFLGRGPLASPGCDHGGFAKVGAGDGDGCDVQFRFLATKSITPDG
 MSTISDSYEAAVDHPDGLTIQTIVARPKSRAGEVKLASRDPAAKPVIENAYLSDEADV
 MTMVKALQKARSASRAPLSAYAGHEEFPGEDVADERQLAAYVRNTAHTANAVVGT
 CKMGESSDALAVVDNHLKVIGVSNLRVVDASVMPTLPGGQTAASTVALAEKAADLIK
 GG

Choline dehydrogenase [*Emiliania huxleyi* CCMP1516]

MSARWLLLLLATHCSAALRNPFRAAPTHFDYIIVGGGTAGCVLADRLSAASKQVLVLEP
 GPSPAAELKIAAPVALTKLFGSEYDWGFRSAPAPGTAGREVHLCRGKCLGGSSATN
 ALLYLRGTAADFQWGLDQWGWSEAMLASFLAVEAQRDAAFRTDALHHGSGGAVPA
 ETPRYANPLSERFLEAAAQAGHPSNADFNDWSRPQAGVGRFQLTTRRGRRAHSAA
 THLRRARRPNLHVRCGCAATRLLEAEGGGGGGGGGGKTRPWTGPAVTGQAGR
 RAVGVEYIDAAGVQRTASVSGGGGGGGGGEVLLCAGAVSSPHLLLLSGIGSPDELA
 HGIGAEVCLPGVGRNLIDQPAVVTGYTVTSPLSITDEMFWRRSGALSPRRVGEWLLR
 GSGPLASSGCDFFGSSRPGLAQPDLQRFVPLGTSPDGVSSYRDIGRAGKTPS
 GLTLQSI AVRPTARGSVSLSSADPSAPPRIETGYGTSEADLATLRQGLRLSRELVAQP
 AFDGVRGEEAWPRAACRLRRPGDDAALDEYIRSTAHSANALGGSCRMGRATSPAR
 LVEGSDPLAVVDPALRVRGASGLRVVDASVLPPLPGGQLGATTFALAERAARIILGER
 AAGEAEAPAERRQEHAHALGAA

Glucose-methanol-choline oxidoreductase [*Chrysochromulina tobini*]

MMRRLVYICAVATVTAAISSRSVPTSARRLIALRGGVAAAEQLAEEPWDYIIVGGGAA
 GCVMAERLSAAEARVLVLEAGTDASRDRLRIRVPAGLIKVKSERDWDFTTEAGQGTS
 GRGIYLCRGKALGGSSCTNVMLYNRGSADYNSWVAAGAEGWGPDSVLHYYRKSE
 NYVGGASQYHGVDGPLSVSDVPYENELSTAFRAAGELGYRRVHDFNDWSAPQEG
 FGRYKVTQRNGERCSAANAYLEGTEGRSNLCVRTGVHATRVTLGSGDDLCAAGV
 EYIGADGKPSRAQLAQQGGEVLLSAGAVQSPQLMLSGIGPRAHLEEVGIEVRKELDN
 VGVGLADHPAVVVSCGSKKKVSVTDEIRLWGGSKTNPALLRWLLWRRGPLTSVAC
 EFGGFFKTKPDLKQADVQVRFVAARAMSPDGITTLQQLGAGAKFLSGYTTQIIACRP
 QSTGLVRLRSSDPLAQMLQDVHLSDDADVATLREGIKLGRQLLAAKSFDQYRDEEV

YPGVAVQSDDEDIDAYVRKTTHSANALVGSCRMGRVDDQAAVLDPEMRVRGVGLR
VVDASAMPHIIGGQTCGPTIMMAEKAADLVLRQRAEINAYMQQAQAYLAASAGAATP
ALSPAQAA

Choline dehydrogenase [*Emiliana huxleyi* CCMP1516]

MVALFALQLALSPQARLGSGSARAALRLRGGSGVTGGSLGRGGGSPAIDGEFDYII
VGGGAAGCVLANRLSADPAHRVLLIEAGGDASRDKRAQVPWAFTKLLRSEYDWDWFH
VEAEAAVNQQEVYLCRGKALGGSSVTNVMLYHRGSPADYDAWEEAGARGWGAKD
VLPYYLRVEDYGDGASQYHAVGGHVSQEVYQNLQSLATFLRAMGQLGFRPNGDF
NDWSSPQEGYGRYKVTQRAGRRTAADGYLAAARERANLVVVTGAQATRLALDSA
YDGAGRLQVSGVEFARGDEREPCSVRLARGGEAVLCAGAVQTPHLLLSGIGPAEH
LREVGVPVRADLPGVGSGLQDHPAVVVSYESKKAVAATDDALLKGYASLVNPLAML
RWLLFGRGPLACAACDHGGFVRRSSPDLQDPDQIRFVPARASSASGMNTLIELGRR
ARFLPGFSTQVVACRPRSEGRVRLRSADPFAKPIIEGIHLGAAEDVASLRHGIRLGRQ
VCAAAAFDEYRGEEVFPGAAVQSDEQIDEYIRSSVHSANALTSSCRMGDPSDPAAVL
DSHLRVRGVGGLRVADASAMPRIIGGQTQAPTYMLAERAADILLHARLQAHEPATES
VSQRLEVAAAAL

Predicted protein, partial [*Phaeodactylum tricornutum* CCAP 1055/1]

YDYIICGGGLAGCVLAERLSQDESKRVLVLEAGGSDYKSLFIRIPAGVLRLFRSKYDW
QHETGGEKGCNGRNVFLQRGKILGGSSCTNVCLHHRGSAEDYNSWNIPGWTATDV
LPFFKQSQKDETGRDATFHGADGEWVMDEVRYQNPLSKLFLEVGEAAGLGTNDDF
NNWSHPQDGVGRFQVSEVNGERCSGATAFLSKAAKRSNVIVRTGTMVRRIDFDETK
TAKGITYDLMGDDTCTVPCLKEGGEVLVTGGAIASPQLLMCSGIGPGKHLRSLGIPVV
HDNSAVGENLQDHPAAVVSFKTPQKGVSVTSKLRFLGKTNPIPVFQWLFFKSGLLTS
TGCDHGAFVRTSDSLEQPDQLQIRFLAARALGPDGMTTYTKFRTMKTVEDGYSFQSV
ACRAKSKGRIRLSSSNHVKPMIDGGYLSNQDDLATLRAGIKLGRMLGNRPEWGEY
LGQEVYPGPDVQTDEEIDEYIRNSLHTANALTGTCKMGTGRGAVVGPDLRVIGVNGV
RVADSSVFPCIPGGQTATPTVMIADRAAVFVR

Probable dehydrogenase [*Chondrus crispus*]

MASPCPAFATPIAVPRSTLTSLISSSSSCTPRPVRT PAPPTHRRLIHMAAPAGTVASTF
RRTVPSSEAATTYDYIIVGGGAAGCVLANRLTEDPSTRVLLLEAGKPDDSFYLVPLG
FPYLLGSPNDWAFVTEPEPNLANRRLYFPRGKVLGGSHAISVMLYHRGHPADYTAW
AESAPGWAPQDVLPHYFLKSESQQSAVFNQDAHGYESPLAVSDLARLNPMASKAFIKA
AHNAAGLNHNPDFNDWATGQDGVGPFQVTQRDGSRESPATSYLRAAKGRRNLTV
MTGAVVERILFENPAGSSTPVATAVSFIDSKGTRVRMSASREILLCGGVYATPQLLML
SGVGPAEHLRSHGIEIVADVPAVGQNLQDHAAAMVSFESQNPEKDKANSSVYYTER
TGKNIGTLLNYVFRGKGPLTSPMCEAGGFAKTDPSMDACDLQRFIPFVSEPDYHS
LADFATAGSYLQNRANRPTGFTIQSVAARPKSRGHVQLRSTDVRDSMSIHGNWISN
DADLKTIVHGVKLCRTIGNDDSMKEFRGRELYPGGEKVSDADIEAYIRDTCHTANAM
VGTCRMIGIEQAAVDPALQVKGVARLRVVDSSVMPTLPGGQSGAPTMMIAEKGADL
IRAAARQADAATVGAAL

Probable dehydrogenase [*Cyanidioschyzon merolae* strain 10D]

MRSRYCFLLSSTPCKYAGQRSFPFASALAGVCAGGRLRNVTNRNLRPGLRTLRSASAE
TEHSQGTRQAQYDFIIVGAGAAGCVLANRLSTAQFSNGDRRYPRVLLLEAGDALAEA

```

PYFEHIPLGFPQLIGSRLDYGFFSRENPTHLGGRGAVYLPRGRGEGGSHAISVMLVH
RGSRDYETWVKDYEALGWGPDDVLPYFKRLESNERTAQRGADGEAATALHGSDG
PLRVSDQRSPNPLSLAFIEACLERGIRRNKDFNDWDHGGQEGAGLFQVTQRDGRRES
PATAYLQPVRSRRNLHIETNALAEHLVWSKDGRRVEGIRFIDRHGRRRAALAHCEVIL
AAGAINTPQLLMLSGLGPGAHLQDFGIPVVRDLPGVGGQNLQDHAAVMLSYYAPDPY
GKDRDKKRIFYTERLGKDPLVLAEYFLLGRGPLTSPVCEAGAFVHTQAVIGEPSCLD
QLRFVPPFFSDADPYKSLGEYRSGGHVLTNTSIRPAGFGLQAVAIRPRSRGRIELATID
PRARPIIHTGWLEDKRDQLTLLSGLKLGREILSGDSMRPYRGREAFPETLEDDLVTYI
RRTCHTANAIVGTARMGTGRDAVVDPELRVHGVERLRVIDASVMPKIIGGQGTGVPTM
MIAERGADLVKKTWKL

```

Plasmids listed in Table 14 used within this work either directly using or as the template for creating further variants.

Table 14. Recombinant plasmids were used in this study.

Plasmid	Description	Reference and source
pET28a(+)	Positive control for lac inducible system	4
pET28a-ANC1-His	Lac inducible expression system with gene. Encoding for ANC1-His expression.	This study
pET28a-sANC1-His	Lac inducible expression system with gene. Encoding for sANC1-His expression.	This study
pET28a-TrxA-sANC1-His	Lac inducible expression system with gene. Encoding for TrxA-sANC1-His expression.	This study
pET28a-ANC2-His	Lac inducible expression system with gene. Encoding for ANC2-His expression.	This study
pET28a-ANC3-His	Lac inducible expression system with gene. Encoding for ANC3-His expression.	This study
pBAD-CvFAP-GFP	L-Arabinose inducible expression system with gene. Encoding for CvFAP-GFP expression.	This study

Recombinant strains containing the plasmids listed in Table 14 were transferred into *E. coli* DH5 α and *E. coli* BL21 (DE3) strains. *E. coli* DH5 α was used for the propagation of plasmids. *E. coli* BL21 (DE3) was used for protein over-expression,

Site-directed mutagenesis was completed with a one-step PCR protocol using the CvFAP or sANC1 plasmids as templates, and the primers are listed in Table 15.

Table 15. Primers were used in this study. Mutated bases are shown in bold letters.

Name	Sequences (5'→3')
T7_pET_mod	CCCGCGAAATTAATACGACTCAC
T7_term	CTAGTTATTGCTCAGCGGT
CvFAP A62G_fdw	CCAATCC GG CTCTGCCGTTGAAGACATC
CvFAP A62G_rev	AACGGCAG AC GCGGACATGGTATATC
CvFAP E152A_fdw	CTTGCG GG CGTCAGATCTACATG
CvFAP E152A_rev	CTGACG CG CCGCAAGCTGTTCCCTG
CvFAP E152N_fdw	CTTGCG AA CCGTCAGATCTACATG
CvFAP E152N_rev	CTGACG GT TCGCAAGCTGTTCCCTG
CvFAP E152G_fdw	CTTGCG GG CCGTCAGATCTACATG
CvFAP E152G_rev	CTGACG GC CCGCAAGCTGTTCCCTG
CvFAP E152S_fdw	CTTGCG AG CCGTCAGATCTACATG
CvFAP E152S_rev	CTGACG G CTCGCAAGCTGTTCCCTG
CvFAP S473G_fdw	CTAAATTCC AGGG CCAGGGTCTGAAATGG
CvFAP S473G_rev	ACCCTGG CC CTGGAATTTAGCAAAC
CvFAP Q474N_fdw	CAGAG CAAC GGTCTGAAATGGCCGAG
CvFAP Q474N_rev	CAGACC AC GCTCTGGAATTTAG
CvFAP E550R_fdw	GATGGT CG CCTGTTCCCAGGTAGC
CvFAP E550R_rev	GAACAG GC ACCATCCAGGTATTCCG
CvFAP E550N_fdw	GATGGT AA CTGTTCCCAGGTAGC
CvFAP E550N_rev	GAACAG G TACCATCCAGGTATTCCG
CvFAP E550G_fdw	GATGGT GG CCTGTTCCCAGGTAGC
CvFAP E550G_rev	GAACAG GC ACCATCCAGGTATTCCG
CvFAP E550K_fdw	GATGGT AA ACTGTTCCCAGGTAGC
CvFAP E550K_rev	GAACAG TTT ACCATCCAGGTATTCCG
CvFAP E550S_fdw	GATGGT AG CCTGTTCCCAGGTAGC
CvFAP E550S_rev	GAACAG G CTACCATCCAGGTATTCCG
sANC1 V366L_fdw	GCTTGT CTG ACCGCTGCGCGTGTT
sANC1 V366L_rev	AGCGGT CAG ACAAGCCGGATGGTC
sANC1 G552S_fdw	CACAG CAG CAACGCTCTGGTCGGT
sANC1 G552S_rev	AGCGTT G CTGCTGTGCACAGTACG
sANC1 T247M_fdw	CGTTG CAT CACCTGGAAGCTC
sANC1 T247M_rev	GGTGAT G CAACGTAAAGGCC

5.2. Preparation of chemo-competent *E. coli* cells

For the preparation of chemo-competent *E. coli* cells, LB-medium (100 mL) was inoculated with an overnight culture of the desired *E. coli* strain. In the case of *E. coli* TOP10 was used, all media were supplemented with kanamycin (40 µg mL⁻¹). All solutions were pre-cooled at 4 °C. Cultures were incubated at 37 °C and 200 rpm until an OD₆₀₀ of 0.4-0.5 was reached. Cells were harvested (4 000 x g, 15 min, 4 °C). The obtained cell pellet was

resuspended in Tfb1 buffer (30 mL) and MgCl₂ solution (1 M, 3.2 mL) and incubated on ice for 15 min. Cells were centrifuged (4 000 x g, 10 min, 4 °C), and the supernatant was discarded before the pellet was resuspended in Tfb2 buffer (4 mL). The suspension was incubated on ice for 15 min before 50 µL aliquots were prepared in 1.5 mL tubes. Aliquots were shock-frozen in liquid nitrogen and stored at -80 °C until use.

5.3. Transformation of competent *E. coli* cells

For transformation, an aliquot of competent cells was thawed on ice, and typically 0.5-1 µL of plasmid DNA or 5-10 µL of a whole-plasmid-amplification reaction (e.g., QuikChange) were added to the cell suspension. After the transformation procedure, cell suspensions were regenerated at 37 °C and 550 rpm for at least 1.5-2 h (Kan resistance). Suitable aliquots were spread on LB agar plates containing antibiotics as required and incubated at 37 °C until colony formation was observed.

5.4. Isolation of plasmid DNA

Cell pellet for plasmid isolation was obtained from 5-10 mL overnight culture by centrifugation or densely grown LB-agar plates. Commercially available GeneJET Plasmid Miniprep Kit (Thermo Scientific) or Wizard Plus SV Minipreps DNA Purification System (Promega) was used to isolate the plasmid DNA according to the manufacturer's instructions. Plasmid DNA was typically eluted in pure ddH₂O.

5.5. Gel extraction and PCR clean-up

For DNA isolation from preparative agarose gels or for clean-up of PCR products, commercially available GeneJET Gel Extraction Kit, GeneJET PCR Purification Kit (Thermo Scientific) Wizard SV Gel, and PCR Clean-Up System (Promega) were used according to the manufacturer's instructions. DNA was typically eluted in pure ddH₂O.

5.6. Colony PCR

For colony PCR, colonies were picked and resuspended in ddH₂O (30 µL). The mixture was heated to 95 °C for 10 min followed by centrifugation, which provided the template for PCR in the supernatant. For amplification, either Red HS Taq Master Mix (Biozym) or OneTaq Quick-Load 2X Master Mix (New England Biolabs) was used. Reaction conditions for colony PCR are summarized in Table 16. The obtained PCR products were analyzed by agarose gel electrophoresis.

Table 16. Sample composition (20 μ L) and temperature program for colony PCR using a ready-to-use 2x PCR Master Mix.

Annealing temperature and elongation time were adapted to the respective reaction setup.

10 μ L	PCR Master Mix 2x
0.4 μ L	Primer fwd [10 μ M]
0.4 μ L	Primer rev [10 μ M]
5 μ L	Colony supernatant
4.2 μ L	ddH ₂ O

95°C	30/60 sec		Initial denaturation
95°C	15 sec	} 30*	Denaturation
Var.	15 sec		Annealing
72°C	var.		Elongation
72°C	5 min		Final elongation
4°C	∞		

5.7. Site-directed mutagenesis

For the introduction of one amino acid exchange, a QuikChange protocol was used. Desired base-pair exchanges were encoded by a pair of primers, which were approx. 20-30 bp long contained the mutation in their 5' complementary region, and both had 3' overhangs, which were only complementary to the backbone. In most cases, a temperature gradient during annealing was chosen according to the calculated melting temperature (T_m) of the respective primers, and elongation time was adapted to the total plasmid size used. A mix containing all reaction components was prepared as in Table 17. Primers are listed in Table 15.

Table 17. Sample composition (20 μ L) and temperature program for QuikChange reaction.

2 μ L	<i>Pfu</i> buffer 10x
0.4 μ L	Primer fwd [10 μ M]
0.4 μ L	Primer rev [10 μ M]
0.1 μ L	<i>Pfu</i> polymerase
0.2 μ L	Template plasmid
0.4 μ L	dNTPs [10 mM]
16.5 μ L	ddH ₂ O

95°C	30/60 sec		Initial denaturation
95°C	15 sec	} 30*	Denaturation
Var.	15 sec		Annealing
72°C	var.		Elongation
72°C	5 min		Final elongation
4°C	∞		

After the reaction, 3 μ L were taken and analyzed on an agarose gel. Meanwhile, 0.5 μ L *DpnI* was added to each sample and incubated at 37 °C for 2 h, followed by heat-inactivation at 80 °C for 20 min. In case of successful amplification, respective samples were used to transform chemo-competent *E. coli* TOP10 cells. Colonies were picked, and the plasmids isolated thereof were sent for Sanger sequencing. After sequencing confirmed incorporation of the desired mutation, competent cells of *E. coli* BL21(DE3) were transformed with the respective plasmids, and transformants were used for heterologous enzyme production.

5.8. Determination of DNA concentration

The concentration of isolated plasmid DNA or purified PCR products was determined spectrophotometrically with a NanoDrop 2000c spectrophotometer (PqLab) at 260 nm. Additionally, the absorption at 280 and 230 nm was measured in parallel to assess the purity of the DNA sample.

5.9. DNA sequencing and sequence analysis

Samples for Sanger sequencing were prepared according to the instructions provided by the company and sent either to GATC Biotech AG or Microsynth Austria GmbH. Respective sequencing primers were either available from the primer collection (Table 15) or provided by the sequencing company. DNA sequences were aligned and analyzed with the online tool Benchling or software SnapGene.

5.10. Agarose gel electrophoresis

Agarose gel electrophoresis was performed to analyze the size of DNA fragments. Depending on the expected fragment size, the prepared gels contained 1-2 % (w/v) agarose (LabQ Agarose LE, LabConsulting) in 1x TAE buffer. The mixture was heated until the agarose was homogeneously dissolved and cooled again to approx. 60°C before DNA-intercalating dye was added (20 000x Green Safe, LabConsulting). The mixture was poured into casts of the desired size, and suitable combs were added. Gels were left until completely solid before used in electrophoresis. Samples meant for agarose gel electrophoresis were combined with 6x DNA Gel Loading Dye (Thermo Scientific) before loading, and GeneRuler DNA Ladder Mix (Thermo Scientific) was used as standard (5 μ L). Analytical gels were routinely run at 120 V and 400 mA for 30-40 min and preparative gels at 100 V and 400 mA for 50-60 min.

5.11. Protein expression

For protein expression, a single colony of freshly transformed cells was first inoculated in 5 mL of liquid LB medium containing 50 $\mu\text{g mL}^{-1}$ kanamycin overnight. The overnight culture was used to inoculate the main culture at 37°C until the desired cell density (OD_{600}) of approximately 0.8 was reached. Overexpression was induced by adding 0.1 mM or 0.5 mM isopropyl β -D-1-thiogalactoside (IPTG); the cultures were grown for another 20 h at 18°C and harvested (4000 rpm, 15 min, 4°C).

5.12. Protein purification

The cells harboring recombinant protein were resuspended in lysis buffer, as shown in Table 22. The cells were disrupted by sonication. The sonicated solution was then centrifuged (12000 rpm, 4 °C, 60 min) to remove any insoluble parts. The soluble fraction was mixed with Ni-NTA (nitrilotriacetic acid) and was incubated at 4 °C for 60 min with low-speed shaking. The column was then washed by gravity flow with the wash buffer (Table 22). The bound protein was eluted with the elution buffer (Table 22). The fractions were collected and analyzed by SDS-PAGE to select the ones containing the target enzyme. The enzyme solution was desalted with the desalting buffer (Table 22) to remove imidazole. Finally, the enzyme samples were concentrated using Amicon® Ultra-15 centrifugal filter device (50 kDa cutoff, Millipore). The final enzyme solution was aliquoted, frozen in liquid nitrogen, and stored at -20 °C.

5.13. Determination of protein concentration

To determine protein concentration, either a Pierce BCA Protein Assay Kit (Thermo Scientific) was used according to the manufacturer's instructions. For calibration, dilutions of bovine serum albumin (BSA) in ddH₂O were prepared in the range of 0.1-2 mg mL⁻¹. Samples were prepared in a microtiter plate and measured on a plate reader (e.g., FLUOstar Omega, BMG Labtech) at 562 nm.

5.14. Preparation of lyophilized cells

The obtained cells were shock-frozen in liquid nitrogen in a suitable container for freeze-drying, and frozen samples were left under reduced pressure (approx. 15 μbar) overnight (AdVantage Pro Lyophilizer, SP Scientific). Lyophilized samples were stored at -20 °C until use.

5.15. Ancestral sequence reconstruction (ASR) photodecarboxylase

In our case, we included 12 sequences that were classified as fatty acid decarboxylases (Table 13). The phylogenetic tree was inferred from included 12 sequences classified as fatty acid decarboxylases by Sorigue and coworkers. Using the online tool GRASP,^[124] we selected different sequence alignment tools and evolutionary models to increase the probability of obtaining an active variant. All algorithms were performed under default

settings. More in detail, ANC1 was generated using T-coffee^[125] for the sequence alignment and the JTT^[126] evolutionary model. ANC2 was generated by changing the evolutionary model to LG^[127], whereas ANC3 was generated using MAFFT^[128] for the sequence alignment and the JTT evolutionary model.

5.16. Homology modeling of AncFAPs

The ancestral AncFAPs were modeled using Rosetta.^[112] The surface residues selection and Images of protein structures were prepared using Pymol.^[129] The hydrophobic scale was computed and represented using ProtScale.^[130] The flexibility comparison was computed and made via MEDUSA.^[113]

5.17. Photocatalytic setup

The photoenzymatic decarboxylation reactions catalyzed by CvFAP and ancestral reconstructed proteins were performed at 30 °C in a total volume of 200 µL Tris-HCl buffer (pH 8.5, 100 mM) containing 30% DMSO or DES as co-solvents. The reaction system was added into a transparent glass vial (total volume 5 mL). The vial was sealed and exposed to blue LED light under gentle magnetic stirring with a speed 200 rpm. At intervals, aliquots were withdrawn, and the reagents were extracted with ethyl acetate (containing 1 mM of cyclohexanol as internal reference) in a 1:1 ratio (v/v). The remaining organic phase was analyzed using GC-FID.

5.18. Melting temperature (T_m) determination using ThermoFAD

Melting temperatures for all enzymes were determined using the ThermoFAD method as first described by *Forneris et al.*^[110] While this method does not assess the unfolding equilibrium, it is valuable for establishing the thermostability of a protein. In a real-time PCR (RT-PCR) machine (Eppendorf) fitted with a 470–543 nm excitation filter and an SYBR Green emission filter (523–543 nm), 20 µl of 1 mg/mL protein were loaded. A temperature gradient from 10 °C to 95 °C was applied (0.5 °C/min), and fluorescence data were recorded every 0.5 °C. A sigmoidal curve was obtained after plotting the fluorescence against the temperature. The unfolding temperature, T_m, is then determined as the maximum of the derivative of this sigmoidal curve.

5.19. Thermostability determination assay with purified enzyme

All enzymatic assays were performed in transparent glass vials sealed with caps having a septum. Reaction mixtures contained 200 µL of each purified enzyme, 5 mM fatty acid (50 mM stock in DMSO), and 30% DMSO as co-solvent. Thermal stability was measured by incubating the purified enzymes at various temperatures for 10 min with occasional shaking. Activities were determined by assaying the residual activity at 30 °C under standard reaction conditions.

5.20. Synthesis of DESs

Depending on the molar ratio between hydrogen-bond acceptor and donor, which is 1:1 or 1:2 as presented in Table 9, the mixtures were prepared by weighing each compound directly into a flask and closing the flasks during the melting for isolating the mixture from air moisture. Then, the eutectic mixtures were created through heating and stirring in the shaker incubator at 50 °C until a transparent liquid was formed.

5.21. Decarboxylation assay in DESs

Purified enzyme/ cell free extract or lyophilized cell was weighted and dissolved in variable amounts of water (within a range of 0–1000 mL). The DES was added to the solution until a total volume of 200 μ L or 500 μ L was reached. Finally, variable amounts of DESs were added to the mixture, which was stirred at room temperature at 300 rpm. After a precise interval of time, the reaction was stopped by adding ethyl acetate, and the mixture was then extracted with ethyl acetate. The organic phases were collected and analyzed with GC-FID.

5.22. Gas chromatography with flame ionization detection (GC-FID)

All chiral GC-FID measurements were performed on a Shimadzu Nexis GC-2010 device with a flame ionization detector (FID) equipped with an AOC-20i/s autosampler and injector unit (Shimadzu) and nitrogen as carrier gas. Methods are summarized in Table 18.

Table 18. Method parameters used for GC-FID analysis.

Item	Value	Units	Control
SPL1 Temperature	230	°C	
SPL1 Pressure	35.8	kPa	
Total Flow	24	mL min ⁻¹	
Purge Flow	3	mL min ⁻¹	On
Primary Pressure	714	kPa	
Column Initial Temperature^[a]	60	°C	
FID1 Temperature	319	°C	
FID1 Makeup^[b]	30	mL min ⁻¹	On
FID1 Hydrogen Flow	40	mL min ⁻¹	On
FID1 Air Flow	399	mL min ⁻¹	On

^[a] Column temperature program: 60 °C for 5 min; 10 °C/min to 200 °C, hold 3 min; 25 °C/min to 300 °C, hold 3 min.

^[b] Carrier gas: N₂.

5.23. In silico design and analysis

Our study selected the generally efficient prediction tools Rosetta software^[101] and HotSpot Wizard^[84] for semi-rational protein engineering.

In Rosetta3.4, the point mutant (pmut) scan application is the computational program that calculates the relative change in folding free energy $\Delta\Delta G$ for predicting protein stability. As shown in Figure 46, the folding free energy for wild-type protein could be calculated with $\Delta\Delta G = \Delta G_3 - \Delta G_2$, $\Delta\Delta G$ is the ΔG for mutant protein minus the ΔG for the wild-type protein.^[131] Therefore, the $\Delta\Delta G$ between the wild-type protein and the mutant can be calculated via $\Delta G_1 - \Delta G_4$. Theoretically, a more stable mutant has a negative $\Delta\Delta G$.

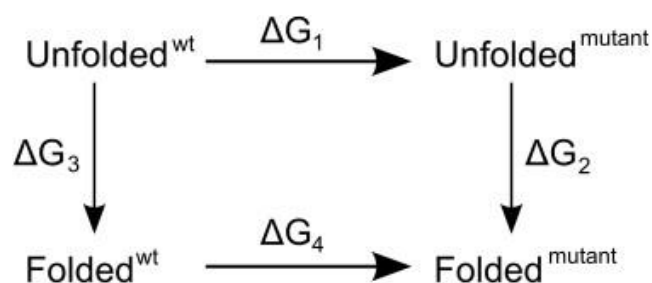


Figure 46. Thermodynamic cycle in Rosetta energy calculation.

In the program, the first step is reading the reference protein structure and scanning all mutants to build a mutants list. The next step is testing and scoring (one Rosetta energy unit or more) of the free energy difference between unfolded and folded protein for all the mutants in the list. The last step is outputting the file that contains the predicted $\Delta\Delta G$.

Another prediction tool we selected for semi-rational protein design to give improved protein stability is HotSpot Wizard,^[84] a web server used to automatically identify hotspots to give desired protein properties such as stability, catalytic activity, substrate specificity, and enantioselectivity.

HotSpot Wizard 3.0 could accept the protein sequence as input data. As shown in Figure 47, the workflow of HotSpot Wizard 3.0 consists of four phases. (1) Constructing of a modeling structure; (2) Annotation of a protein; (3) Identification of mutagenesis hot spots; (4) Design of mutations and a smart library.

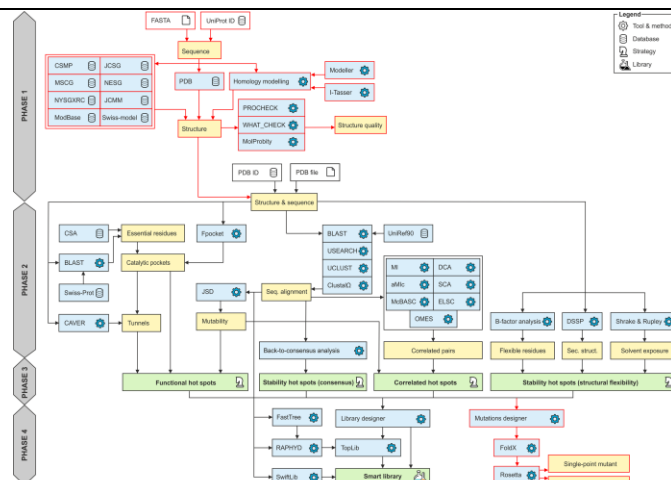


Figure 47. Workflow of HotSpot Wizard 3.0.

Initially, we selected entering the PDB ID CvFAP (5NCC) as the template for modeling, then the process of hotspot identification begins automatically after entering the template structure.

chain	position	residue	secondary stru.	relative ASA (%)	B-factor rank	relative flexibility	HotSpot
A	82	Ala	loop	81	3	high	✓
A	83	Ser	loop	68	7	high	✓
A	86	Ile	alpha helix	89	8	high	✓
A	89	Arg	loop	83	8	high	✓
A	70	Lys	hydrogen bond	81	9	high	✓
A	409	Arg	extended strand	25	10	high	✓
A	410	Gln	extended strand	48	2	high	✓
A	475	Gln	extended strand	75	6	high	✓
A	479	Gly	extended strand	22	1	high	✓
A	477	Leu	extended strand	88	4	high	✓

Figure 48. The interface of the prediction result output in the HotSpotWizard 3.0.

5.24. Nile Red staining and fluorescence spectrometry

Nile red from the Aldrich Chemical (Milwaukee, USA) was dissolved in 0.1 mg/ml with DMSO (Dimethylsulfoxide). Fluorescence spectra were measured with the spectrofluorometer FP-750 with PC control equipment. The acquiring and processing of the data were done using the PC software. Nile red fluorescence was measured at 450 nm excitation with emission from 500 nm to 700 nm using a Tecan M1000 microplate reader.

The cells were harvested from 5 mL of overnight culture by centrifugation at 4000 rpm for 10 min using a 1.5 mL centrifuge tube. Decanting the supernatant and allowing the pellet

to drain, removing as much supernatant as possible. The cell lysis was achieved by thawing and resuspending at room temperature in BugBuster reagent to obtain cell free extract. Incubate the cell suspension on the rotating mixer at a slow setting for 20 min at room temperature. The solution was then centrifuged (12000 rpm, 4 °C, 20 min) to remove insoluble cell debris.

Next, 100 µL of cleared cell free extracts were added to a clear-bottomed 96-well polystyrene plate containing 30% DMSO and Tris-HCl (pH 8.5, 100 mM), the plate was incubated at room temperature for 10 minutes. As showed in Table 2, each condition was initiated by adding different concentrations of contents, respectively. The emission spectra from 500 nm to 700 nm of each well were recorded at 450 nm excitation Tecan M1000 microplate reader.

Table 19. Chemicals used in this study.

Substrate	CAS No	MW
Caprylic acid C8	124-07-2	144.21
Capric acid C10	334-48-5	172.26
Heptane C7	142-82-5	100.20
Nonane C9	111-84-2	128.26
Decane C10	124-18-5	142.28
Dodecane C12	112-40-3	170.33
Undecane C11	1120-21-4	156.31
Dodecanoic acid C12:0	143-07-7	200.32
Tridecane C13	629-50-5	184.36
Myristic acid C14	544-63-8	228.37
Pentadecane C15	629-62-9	212.41
Palmitic acid C16	57-10-3	256.42
Stearic acid C18:0	57-11-4	284.48
Oleic acid C18:1	112-80-1	282.46
Linoleic acid C18:2	60-33-3	280.45
DL-Menthol	89-78-1	156.27
Tetra-n-butylammonium chloride hydrate	37451-68-6	277.92

Table 20. Composition of general media, buffers, and solutions used in this study.

LB medium	
Tryptone	10 g
Yeast extract	5 g
NaCl	10 g
dH ₂ O	To 1 L and autoclave.
10x TB salts	
KH ₂ PO ₄	2.31 g
K ₂ HPO ₄	12.54 g

dH ₂ O	To 100 mL and autoclave.
TB medium	
Tryptone	12 g
Yeast extract	24 g
Glycerol	5 g
dH ₂ O	To 900 mL and autoclave. Combine with 100 mL 10x TB salts before use.
Kanamycin 1000x (Kan)	
Dissolve in ddH ₂ O and sterilize by filtration (0.22 µm). Store aliquots at -20 °C.	
Kanamycin sulphate	40 mg mL ⁻¹
Ampicillin 1000x (Amp)	
Dissolve in ddH ₂ O and sterilize by filtration (0.22 µm). Store aliquots at -20 °C.	
Ampicillin sodium salt	100 mg mL ⁻¹
Continuation of Table 22.	
IPTG (1 M)	
Dissolve in ddH ₂ O and sterilize by filtration (0.22 µm). Store aliquots at -20 °C.	
Isopropyl-β-D-thiogalactopyranoside (IPTG)	238.3 mg mL ⁻¹
50x TAE (pH 8) Add ddH ₂ O to 1 L and adjust pH to 8.	
Tris	242.3 g
EDTA	14.6 g
Acetic acid	57 mL

Table 21. Buffers and solutions for chemo-competent *E.coli* cells preparation in this study.

TfB1 buffer	
CH ₃ COOK	1.47 g (0.03 M)
KCl	3.73 g (0.1 M)
CaCl ₂	0.55 g (0.01 M)
Glycerol	65 mL (13 %)
dH ₂ O	To 500 mL and autoclave.
TfB2 buffer	
KCl	0.37 g (0.01 M)
CaCl ₂	5.55 g (0.1 M)
MOPS	1.05 g (0.1 M)
Glycerol	65 mL (13 %)
dH ₂ O	To 500 mL and autoclave.
MgCl₂ solution Prepare the solution in dH ₂ O and sterilize by filtration.	
MgCl ₂	0.0952 g/mL (1 M)

Table 22. Buffers for enzyme purification in this study.

Lysis Buffer	
---------------------	--

NaH ₂ PO ₄	6.24 g
NaCl	17.53 g
Glycerol	100 mL(10%)
Imidazole	0.68 g
dH ₂ O	To 1000 mL, adjust to pH 8.0 with NaOH
Wash Buffer	
NaH ₂ PO ₄	6.24 g
NaCl	17.53 g
Glycerol	100 mL(10%)
Imidazole	1.36 g
dH ₂ O	To 1000 mL, adjust to pH 8.0 with NaOH
Elution Buffer	
NaH ₂ PO ₄	6.24 g
NaCl	17.53 g
Continuation of Table 22 .	
Glycerol	100 mL(10%)
Imidazole	17.02 g
dH ₂ O	To 1000 mL, adjust to pH 8.0 with NaOH
Desalting Buffer	
NaH ₂ PO ₄	6.24 g
Glycerol	100 mL(10%)
dH ₂ O	To 1000 mL, adjust to pH 7.4 with NaOH

Table 23. Buffers and solutions for SDS-PAGE in this study.

Tris buffer A	
Tris	45.4 g
SDS	1 g
dH ₂ O	To 250 mL, adjust to pH 8.8 with HCl
Tris buffer B	
Tris	15.1 g
SDS	0.25 g
dH ₂ O	To 250 mL, adjust to pH 6.8 with HCl
10x Running buffer	
Tris	30 g
Glycin	144 g
SDS	10 g
dH ₂ O	To 1 L, adjust to pH 8.3 with HCl
4x Sample buffer	
Tris-HCl 0.5 M, pH 6.8	1.25 mL
10% (w/v) SDS in dH ₂ O	2 mL (can be stored at r.t.)
0.5% (w/v) bromphenol blue in dH ₂ O	0.2 mL (can be stored at -20 °C)
Glycerol	2.5 mL
dH ₂ O	3.55 mL

Prepare 950 μ L aliquots and store at -20 °C. Add 50 μ L β -mercaptoethanol per aliquot before first use.

Staining solution

Coomassie Brilliant Blue R-250	1 g
Acetic acid, conc.	100 mL
Ethanol, 100%	300 mL
dH ₂ O	600 mL

Dissolve Coomassie in ethanol first, add water and acetic acid last. Filtrate before first use.

Destaining solution

Acetic acid, conc.	100 mL
Ethanol, 100%	300 mL
dH ₂ O	600 mL

The solution can be regenerated over charcoal after use.

CHAPTER II

Discovery and characterization of a Baeyer-Villiger Monooxygenase (BVMO) from *Cupriavidus necator* H16.

1. Abstract

Baeyer-Villiger monooxygenases (BVMOs) are a well-known and systematically studied class of flavin-dependent enzymes. They are biocounterpart to chemical catalysts of Baeyer-Villiger (BV) oxidations in synthetic chemistry. The large diversity of BVMOs enables oxidization of numerous substrate types (e.g., linear, cyclic, and aromatic ketones, steroids, and terpenoids). Moreover, the high chemo-, region- and enantioselectivity of BVMOs provide access to the synthesis of valuable compounds, which are impossible to achieve by chemical approaches.

However, despite extensive study of BVMOs, the industrial application part is still missing. The limitation of BVMOs industrial application is their low stability and cofactor NAD(P)H dependency. Currently, the investigations about either for immobilization of BVMO expressing whole-cell or isolated enzymes was the attempt to improve the overall biocatalyst stability. The investigated solution was either adding a second enzyme for cofactor regeneration or applying the whole cell system to overcome the challenge of cofactor dependency. Nevertheless, all these strategies were only enabled to tackle one limitation at once, and the processes were complicated.

As our results presented in this chapter, we identified 18 putative BVMOs encoding genes from the genome of chemolithoautotrophic bacterium *Cupriavidus necator* H16. *C.necator* H16 comprises the efficient NADH regeneration dehydrogenases, which provide the cofactor NADH for BVMOs.

After gene screening via Pattern Hit Initiated BLAST (PHI-BLAST) with consensus sequences of BVMOs, we identified 18 putative BVMOs encoding genes from the genome of *C.necator* H16. After sequence analysis on nucleotide level, three flavin-dependent monooxygenases containing typical sequence motifs for type Type (I) BVMOs (FXGXXXHXXXW and two GXGXXG-motifs) and five hypothesized Type "O" BVMOs (two GXGXXG-motifs, without activity related motifs) were identified.

The study was focused on one putative BVMOs (named CPMO0) firstly. A set of studies to heterologous express CPMO0 in *E.coli* BL21(DE3) was performed. However, the enzymes overexpressed in *E.coli* were inactive and insoluble. Latter, the soluble active enzymes were obtained from *C.necator* H16. Characterization of CPMO0 was attempted. CPMO0 could convert alicyclic ketone (Cyclopentanone), bicyclic ketones (Bicyclo[3.2.0]hept-2-en-6-one), and sulfides (Thioanisole and Dimethylsulfoxide). From the thermoFAD study, this enzyme showed a high thermostability in various conditions.

On the one hand, our results were the first study about BVMOs from *C.necator* H16, and we determined one novel active BVMO, further study about the rest of putative BVMOs would provide more interesting results. On the other hand, *C.necator* H16 was a suitable microbial platform to produce protein and chemicals on a large scale, which would favor the industrial applications of BVMOs.

2. Introduction

2.1. Oxygenases

In microbial catabolic pathways, oxygenases play significant roles.^[132] For example, they enable the activation of C-H bonds for alkane degradation.^[133] They initiate the degradation of aromatic compounds by hydroxylating the aromatic ring to prepare for ring cleavage and catalyze the ring fission reaction.^[134] Oxygenases form an essential class of enzymes applied for the biocatalytic oxidation of several compounds, then eventual transformation of these raw materials into valuable products or chemicals.^{[135],[136]} Comparing with the conventional chemical catalytic oxidation reactions, oxygenases-mediated reaction exhibit few advantages: 1, reaction readily achieved or controlled; 2, the chemical counterparts is not regioselective or stereoselective; 3, oxygen is the cheaper and more environmentally friendly oxidant comparing with toxic chemical oxidants; 4, Introduction of chirality.^[137]

Oxygenases can be grouped into two categories, i.e., mono- and dioxygenases (Figure 49). Monooxygenases incorporate one atom from dioxygen into substrates.^{[138],[139]} Dioxygenases introduce both oxygen atoms into substrates.^[140] In order to insert one or two oxygen atoms into a substrate molecule, oxygenases usually need NAD(P)H as an electron-donating coenzyme, while metal or organic cofactor is required for oxygen functionalization.^[140]

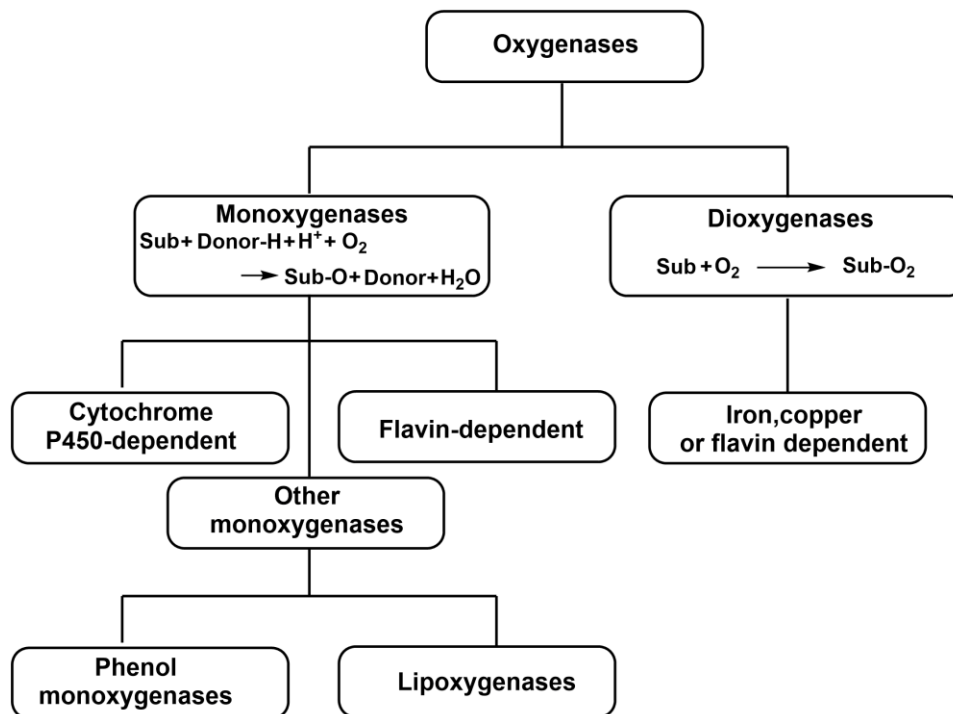


Figure 49. Classification of oxygenases and enzymatic oxidation reactions as catalyzed by monooxygenases and dioxygenases.

2.2. Flavin-dependent monooxygenases

A wide variety of biological processes involves flavin-dependent monooxygenases (FMOs), ranging from siderophore^[141] and auxin biosynthesis,^[142] lignin and xenobiotics degradation,^{[143],[144]} antibiotics,^[145] antitumor,^[146] to light emission.^[147] Flavin monooxygenases perform oxygenation reactions, i.e., hydroxylations, Baeyer-Villiger oxidations, sulfoxidation, epoxidations, and hydrogenation. Therefore, FMOs represented highly attractive biocatalytic tools with their capacity to catalysis various monooxygenation reactions with outstanding selectivity. Many of the monooxygenation reactions are difficult or even impossible to achieve with chemical methods.^[148]

Depending on fold and function, flavin monooxygenases can be divided into six subclasses. (Table 24) Subclasses A and B protein utilize NAD(P)H as the external electron donor. Flavin reduction and oxygenation happen in the single polypeptide chain. Subclasses C-E are two-component enzymes composed of a NAD(P)H-dependent flavin reductase and a flavin-specific oxygenase.

Table 24. Flavin monooxygenase classification.^[149]

Subclass	Structural fold	Subunits	Cofactor	Coenzyme	Reaction
A	Glutathione reductase-2	α	FAD	NAD(P)H	Hydroxylation
B	NADP Rossmann	α	FAD	NADPH	Baeyer-Villiger oxidation N, S, P, Se, I- oxygenation N-hydroxylation
C	Tim Barrel	$\alpha+\beta$	-	FMN NAD(P)H	Light emission Baeyer-Villiger oxidation Desulfurization, sulfoxidation Hydroxylation
D	Acyl-CoA dehydrogenase	$\alpha+\beta$	-	FAD NAD(P)H	Hydroxylation
E	Glutathione reductase-2	$\alpha+\beta$	-	FAD NAD(P)H	Styrene epoxidation
F	Glutathione reductase-2	$\alpha+\beta$	-	FAD NAD(P)H	Halogenation

2.3. Baeyer-Villiger monooxygenases (BVMOs, E.C. 1.14.13.xx)

In 1899, Adolf von Baeyer (Nobel laureate in chemistry for the synthesis of indigo in 1905) and his student Victor Villiger reported the first time for utilizing potassium monopersulfate (KHSO_5 ; also known as Caro's acid) as the oxidant to transform ketones into esters or cyclic ketones into lactones.^[150] Hence, this novel reaction had been named Baeyer-Villiger (BV) oxidation.

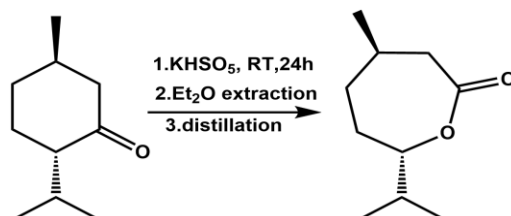


Figure 50. Oxidation of menthone by Baeyer and Villiger. Figure adapted from Leisch.^[151]

Now more than a century, this famous Baeyer-Villiger (BV) oxidation transformation has been developed for wide-ranging applications, from the synthesis of steroids, antibiotics, and pheromones to the synthesis of monomers for polymerization.^[152] Baeyer-Villiger (BV) oxidation is frequently employed in synthetic organic chemistry, but the disadvantage of chemical synthesis lacks the selectivity of chemical catalysts and the harsh conditions required for reactions. Therefore, the enzymatic counterparts are attractive to use as biocatalysts.

In 1948, Turfitt provided the first indication of the biological BV reaction in a series of microbiological investigations of steroids.^[153] Trudgill and co-workers provided the first Baeyer-Villiger monooxygenase (EC 1.14.13.x). Their study showed this novel enzyme origin from the *Acinetobacter* bacterium. It was a flavoprotein with a broad ketone substrate acceptance profile.^[154]

Latter, more BVMOs have been identified and studied. It is a class of NAD(P)H and FAD or FMN-dependent enzymes. BVMOs are capable insert an oxygen atom from aerial oxygen into a C-C bond of ketones, cyclic ketones, and aldehydes. As well as enable to catalyze the oxidations of heteroatoms (*i.e.*, sulfides, amines, and boron compounds).^{[155]-[157]} In addition, these enzymes also enable the oxidation of substrates with very high chemo-, regio- and enantioselectivity.^[158]

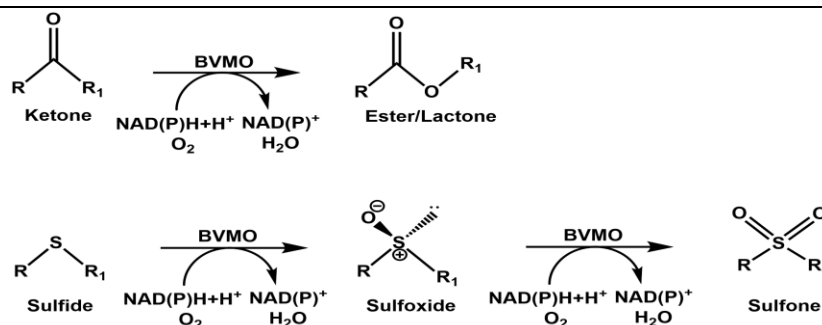


Figure 51. Baeyer–Villiger oxidation and sulfoxidation reactions catalyzed by flavin-dependent Baeyer–Villiger monooxygenases (BVMOs).

Up to now, BVMOs were systematically studied and described in several recent articles.^{[151],[158],[159]} A few studies have focused on revealing the enzymatic mechanism of BVMO. The illustrative example is CHMO from *Rhodococcus* sp. HI-31, during the catalytic cycle, several crystal structures presented the movement of enzyme domains.^{[160]–[162]} The structure study of BVMO phenylacetone monooxygenase (PAMO) from *Thermobifida fusca* provided the information about rationally designing of active-site to improve enzyme properties, the catalytic arginine residue R337 was highlighted as key residue in mechanism and catalysis.^[163] Other studies demonstrated the high versatility of BVMO mediated reaction. MtmOIV from *Streptomyces argillaceus* was found as the key oxygenase involved in mithramycin biosynthesis and may also catalyze various following reaction steps to synthesis numerous unknown intermediates.^{[164],[165]} The BVMO FR901464 from *Pseudomonas* sp. no. 2663 enables operating in cis within the PKS and NRPS biosynthetic paradigm.^[166] PtIE from *Streptomyces avermitilis* involved in the biosynthesis of neopentalenolactone.^[167]

The fast-growing information deposited in public sequence databases favored the identification of numerous new BVMOs. The present classification of BVMOs is Type I, II, and atypical Type “O”.^[168] General characteristics of the respective types as shown in Table 25. The most recombinantly available BVMOs are the Type I BVMOs, which are NADPH and FAD-dependent. As Fraaije *et al.* demonstrated, Type I BVMOs consist of some consensus motifs such as FXGXXXHXXXW, which are the functional motifs for enzyme catalysis and substrate specificity. Two GxGxx(G/A) Rossmann fold motifs enable bind to the pyrophosphate moiety of FAD and the NAD(P)H cofactor. These fingerprint motifs facilitated the identification of putative BVMOs from sequence data.^[169]

Table 25. General Characteristics and Classification of BVMOs.

Type (I) BVMO	Type (II) BVMO	Type “O” BVMO
Flavin-dependent monooxygenase subclass B	Flavin-dependent monooxygenase subclass C	Flavin-dependent monooxygenase subclass A

Single gene product	Two separate gene products, reductase and oxygenase	Single gene product
Contain a tightly bound FAD cofactor	Use reduced FMN generated by the reductase as a coenzyme	FAD bound to the surface of the protein
NADPH-dependent NADPH/NADP ⁺ bound during catalysis	The reductase can use NAD(P)H	NAD(P)H-dependent
“BVMO” fingerprint motif sequence: FXGXXHXWP Two domain structures resembling disulfide oxidoreductases; two dinucleotide-binding domains (Rossmann fold) binding FAD	Absence of “BVMO” fingerprint The structural core of the oxygenase subunit displays a TIM-barrel fold	Absence of “BVMO” fingerprint absence of a structural relationship to FAD-dependent hydroxylases

Few examples of the enzyme by gene mining with the fingerprint motifs have been identified as BVMOs. The newly identified 116 BVMO sequences from Universal Protein Resource database.^[170] The Grogan group discovered 23 putative BVMO genes in the genome of *Rhodococcus jostii*.^[171] Singh *et al.* isolated a BVMO from the metagenomics library of an effluent treatment plant.^[172] The increasing number of novel BVMO appearing within the gene databases indicates a rich reservoir of BVMOs with promising unknown activity and selectivity attributed to natural evolutionary processes.^[173]

Up to now, a large number of BVMOs have been discovered, cloned, and characterized by different microorganisms. As shown in Figure 52, the associations of BVMOs exist among largely divergent species. One of the hypothesized explanations for the lineages that BVMOs may be essential (*e.g.*, in defense metabolite synthesis) to preventing gene loss during the evolution of these species, such as the significant number of BVMO encoding genes of the genus *Aspergillus* or *Rhodococcus*.^{[171],[174],[175]}

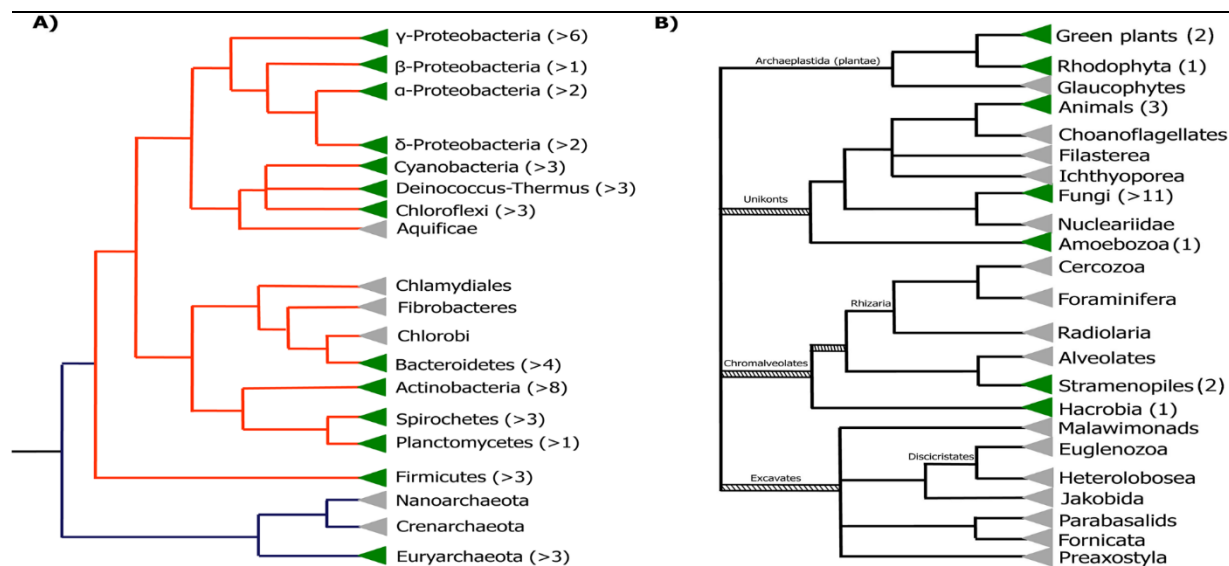


Figure 52. Taxonomic distribution of BVMO genes.^[176]

A. The distribution of BVMO encoding genes in genomes from Bacteria and Archaea domains. B. The distribution of BVMO encoding genes in genomes from the Eukarya domain. Green triangles show clades containing BVMOs; grey triangles show clades where BVMOs were not detected. Numbers in brackets show the number of different species where BVMOs were found. The tree topology is based on the tree of life web project (Maddison DR and Schulz KS (eds.) 2007. Available at: <http://tolweb.org> (Accessed 17 November 2014)).

Due to the substrates promiscuity, high chemo-, regio- and enantioselectivity and abundantly presence in nature, BVMOs are the highly attractive biocatalyst for synthetic chemistry. Nonetheless, only a minimal number of BVMOs industrial applications were demonstrated.^{[177],[178]} Few drawbacks for BVMOs were reported, such as the low stability and activity, cofactor NAD(P)H dependency, the inhibition by substrates and products, adequate oxygen transfer, and tolerance of organic cosolvents.^{[179],[180]} Currently, the investigations about either for immobilization of BVMO expressing whole-cell or isolated enzymes was the attempt to improve the overall biocatalyst stability.^{[181],[182]} For the challenge of cofactor dependency, the investigated solution was either adding a second enzyme for cofactor regeneration or applying the whole cell system.^[183] However, all these strategies were only enabled to tackle one limitation at once, and the processes were complicated.

2.4. *Cupriavidus necator* H16

C. necator H16 is a gram-negative, facultatively chemolithoautotrophic and soil-dwelling bacterium of the class β -proteobacteria.^[184] It was isolated from soil in Germany^[185] and belonged to the most widely investigated hydrogen-oxidizing *Knallgas* bacteria capable of growing at the interface of anaerobic and aerobic environments.^[186]

The genome of *C. necator* H16 was sequenced and annotated in detail in 2006 (Gen Bank accession number: ASM928v2).^[187] The total size of *C. necator* H16 genome is 7,416,678 bp and consisted by three circular replications: chromosome 1 (4,052,032 bp), chromosome 2 (2,912,490 bp), and megaplasmid pHG1 (452,156 bp).^[188] The general features of the three replications are presented in Table 26. The GC content of two chromosomes almost are identical, the proportion of coding regions are nearly the same. The two parameters for megaplasmid pGH1 differ from two chromosomes. There are 59 transfer RNA (tRNA) genes in *C. necator* H16, 51 complete sets representing all possible codons are carried on chromosome 1, 7 duplicates are found on chromosome 2, and only one is located on pHG1. Three complete ribosomal RNA (rRNA) operons were identified on chromosome 1 and two on chromosome 2.

Interestingly the most of the putative BVMO encoding genes were located at chromosome 1. The major functional genes distributed on chromosome 1 encodes most key enzymes for DNA replication, transcription, and translation. The genes on chromosome 2 for central steps of sugar and sugar-acid metabolic path, the degradation of aromatic compounds, and utilization for different nitrogen sources. All three replicons harbor genes for encoding protein involved in the synthesis of cell appendices and chemotaxis.^{[187],[189],[190]}

Table 26. General features of the *Cupriavidus necator* H16 genome

Feature	Chromosome 1	Chromosome 2	Megaplasmid pHG1
Size (bp)	4,052,032	2,912,490	2,912,490
G+C ratio (mol%)	66.4	66.7	62.3
Percentage coding	88.1	88.6	79.7
tRNA	51	7	1
rRNA operons	3	2	0
IS elements	4	3	7
Total number of CDSs	3651	2555	420
No. of CDSs with assigned function	2382	1618	225
CDSs with unknown function	841	680	157
CDSs with general function prediction	428	257	32

C.necator H16 exhibits a very flexible metabolism, and it could grow heterotrophically, autotrophically, or in a mixotrophic fashion. It accepts diversified growth substrates under heterotrophic growth conditions, including several organic acids such as acetic acid and sugars like fructose.^[185] *C.necator* H16 enable to metabolize these substrates via the Entner-Doudoroff (ED) pathway and the tricarboxylic acid (TCA) cycle.^{[187],[190]}

Under lithoautotrophic growth conditions, CO₂ and H₂ can serve as the sole carbon and energy sources, respectively.^[191] CO₂ is assimilated through the Calvin-Benson-Bassham (CBB) cycle, and H₂ be oxidized by hydrogenases (SH).^[192] The NAD-reducing [NiFe]-hydrogenase [SH; EC 1.12.1.2] has been reported as a high efficient NADH regeneration catalyst.^[193] The CPCR (carbonyl reductase from *Candida parapsilosis*)-mediated reduction of acetophenone in the coupled systems employing either SH or FDH (the formate dehydrogenase from *Candida boidinii*), which as the NADH regeneration catalysts. The turnover number for SH was 143,666 and 1685 for SH and FDH.^[194] Under organoautotrophical growth conditions, *C.necator* H16 is capable oxidize such as formic acid to CO₂ and generate NADH by a membrane dehydrogenase (MBH).^[195]

The dependency for expensive cofactor, i.e., NAD(P)H, is the hurdle for BVMOs industrial applications. Since biocatalytic systems employing whole cells can be used, which circumvent expensive coenzyme regeneration procedures.^{[168],[196]} Using *C. necator* H16 whole-cell as biocatalyst would overcome the two challenges for industrial applications of BVMOs.

Besides the outstanding ability to accept the great diversity of carbon and energy sources, *C.necator* can grow to high-cell densities under heterotrophic and lithoautotrophic conditions.^{[197],[198]} Moreover, unlike other bacteria, *C.necator* H16 can be cultivated in high-cell density fermentations without the accumulation of inhibiting organic acids. Therefore, combining with the flexible metabolism along with its natural ability to grow high cell densities,^{[187],[199],[200]} *C. necator* H16 can be recognized as the ideal candidate for producing value-added compounds (e.g., fuels,^[191] polymers,^[201] and fine chemicals^[202]) with the fermentation process.

3. Results and discussion

3.1. Discovery of Baeyer-Villiger monooxygenases from *C.necator* H16.

For this study, gene screening via Pattern Hit Initiated BLAST (PHI-BLAST) with fingerprint motifs of BVMOs (Type I BVMOs comprise FXGXXXHXXXW and two GXGXXG-motifs; Type “O” BVMOs comprise two GXGXXG-motifs) was used to identify putative BVMOs encoding gene in *C.necator* H16 genome.

As listed in Table 27, eighteen putative genes encoding BVMO were identified from the *C.necator* H16 genome. Screening to 3806 genes (Table 42) locating at chromosome 1, thirteen genes were identified as BVMOs. Screening with 2603 genes (Table 43) locating at chromosome 2, four genes were identified as BVMOs. One BVMO gene from 428 genes (Table 44) locates at megaplasmid pHG1. All 18 putative BVMOs sequences are highly divergent, and the identity between sequence pairs ranges from 10% to 31%.

Table 27. Overview of putative BVMOs encoding gene from *C.necator* H16

No.	Gene ID	Location	Gene symbol	Size (bp)	Protein Accession Number
0	57645873	Chromosome 1	E6A55_RS19085	1575	WP_011616290.1
1	57645129	Chromosome 1	E6A55_RS15325	882	WP_010814951.1
2	57645015	Chromosome 1	E6A55_RS14755	1305	WP_011615888.1
3	57644864	Chromosome 1	E6A55_RS14000	1326	WP_011615806.1
4	57644853	Chromosome 1	E6A55_RS13945	1266	WP_011615802.1
5	57644232	Chromosome 1	E6A55_RS10810	1341	WP_011615498.1
6	57643928	Chromosome 1	E6A55_RS09265	1428	WP_011615329.1
7	57643691	Chromosome 1	E6A55_RS08035	1845	WP_011615168.1
8	57643685	Chromosome 1	E6A55_RS08005	1476	WP_011615163.1
9	57643417	Chromosome 1	E6A55_RS06665	1107	WP_011614996.1
10	57643297	Chromosome 1	E6A55_RS06050	1086	WP_010808985.1
11	57643242	Chromosome 1	E6A55_RS05775	1569	WP_011614904.1
12	57642392	Chromosome 1	E6A55_RS01505	1128	WP_010814867.1
13	57647647	Chromosome 2	E6A55_RS28025	2020	WP_011617449.1
14	57648403	Chromosome 2	E6A55_RS31855	1227	WP_011617926.1
15	57646843	Chromosome 2	E6A55_RS23985	1971	WP_011616905.1
16	57646666	Chromosome 2	E6A55_RS23090	1242	WP_010814240.1
17	39976490	Megaplasmid pHG1	E6A55_RS34370	1380	WP_011154326.1

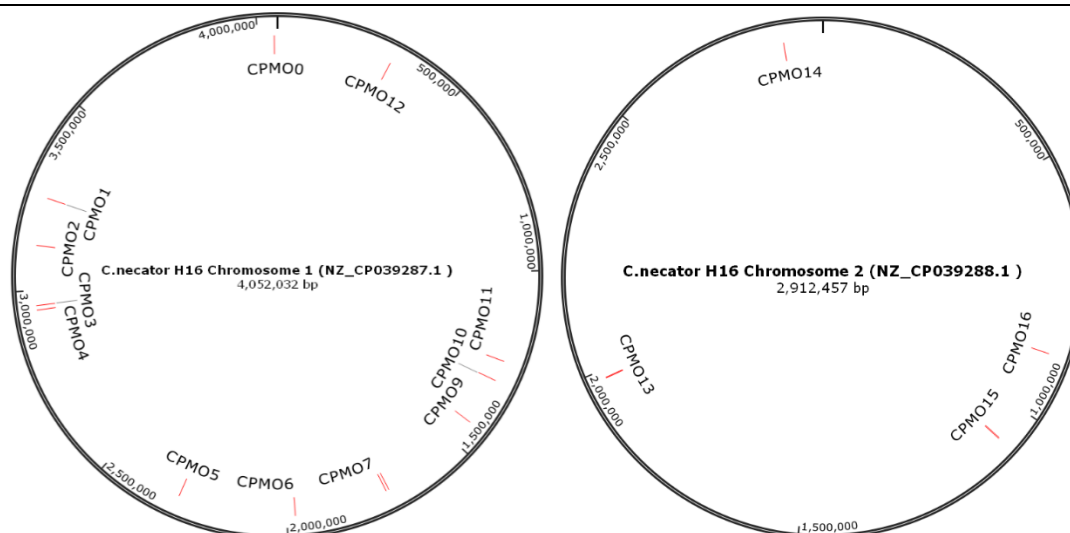


Figure 53. The distribution of the identified BVMO genes on *C. necator* H16 chromosome 1 and 2.

Multiple nucleotide sequence alignment was performed between these putative BVMO sequences and other typical BVMOs. Four characterized BVMOs were selected: 1, CHMO is Cyclohexanone monooxygenase from *Acinetobacter calcoaceticus*.^[203] 2, HAPMO is 4-hydroxy acetophenone monooxygenase from *Pseudomonas fluorescens*.^[204] 3, CtCPMO is Cyclopentanone monooxygenase from *Comamonas testosteroni*.^[205] 4, PAMO is from *Thermonifida fusca*.^[206]

As listed in Table 28, five of the sequences consist of two Rossmann fold motifs (GxGxxG/A), indicating they code for flavin monooxygenases. It is worthy of mentioning, as the atypical BVMOs, type 'O' BVMO does not contain typical consensus sequences FxGxxxHxxxW but contains two Rossmann fold motifs (GxGxxG/A).^[164] Further discrimination for the presence of the "fingerprint" sequence motif (FxGxxxHxxxW)^[169] resulted in three sequences (named CPMO0, CPMO11, and CPMO13) were envisioned as BVMOs. The overview of multiple sequence alignment of CPMO0, CPMO11, and CPMO13 with four known BVMOs was presented in Figure 54.

Table 28. Sequence motifs of putative BVMOs from *C.necator* H16.

	Gene ID	Rossmann fold motifs (GxGxxG/A)	Type 1 BVMO consensus FxGxxxHxxxW	Rossmann fold motifs (GxGxxG/A)
CPMO0	57645873	GSGFAG	FEGECFHSQHW	GTGASA
CPMO1	57645129	GLGAMG		GAAPGT
CPMO2	57645015	GAGFTG	ICGKTVYRIVW	
CPMO3	57644864	GAGGLE	LIAACVRAQNW	GAGATG
CPMO4	57644853	GAGAAG	FLLLTATGMVR	IPKIGA
CPMO5	57644232	GAGPYG	PTYVVSSEF	GGQSA
CPMO6	57643928	GGGFAG		GGFSG
CPMO7	57643691	GGGAG	GTTDYLPLSG	LRDKAG
CPMO8	57643685	GGMAG	GLVALSMQTPD	RDGLGG
CPMO9	57643417	GAGVVG		DSHGLM
CPMO10	57643297	GAGPVG	LEGRHVHYAVR	GGDSA
CPMO11	57643242	GAGLSG	FRGSMVHPQFW	GSGATA
CPMO12	57642392	GASFAG	FVGAVPRNGGW	GAKSRV
CPMO13	57647647	GAGMSG	FKGVSIHANRW	GTGATG
CPMO14	57648403	GAGPAG	VSGAQPEVTVR	LLDASG
CPMO15	57646843	GTGPVG	GRGIRVSDRMQ	REQVYG
CPMO16	57646666	GNGMAG	VIAFRDIQDVE	GGLLG
CPMO17	39976490	GGGAGG	IRAHAQPAPLE	GAGATD

CHMO	-----MTAQMKTNASAKKAVNSPNHIFDTFIVGAGTSGIATAIRL	36
CPMO0	-----MNAPEALPQQASPALEHTDAITIGSGFAGLGMAIRL	40
HAPMO	MDLETEAQQVREFCLQLRIDFRDSGQPAPGRPTSDQLHILGTLWLMGVPVIEPYLPLTAAEEAVTAEEDLRAPRWHKDHVASGRDFKVVIIIGAGESGLYIALRL	100
CtCPMO	-----MTTMTTTEQLGMNNSVNDKLDVLLIGAGFTGLYLR	40
PAMO	-----MAGQTTVDSRRQPP-----EEVDVLVVGAGFSGLYALYRL	35
CHMO	DQVGVTNYKIIIEKASRVGGTWTRENTYPGCGDVPVLSYYS----FAPSAKWLSHLFARQPELTVLEDSVNEFNVASKIEFNNELLNAAWDSRHLWVLD	136
CPMO0	RERGERNFLIFEKAGSVGGTWRDNHYPGCACDQVSHLSYFS----FAPNPDWSRMSFTQPEIRAYLEGGCTDRFGLRPHRLRHLHELRAAWDEAAQVWRLE	132
HAPMO	KQAG-VPFVVEYKGNVDVGGTWTRENTYPGCRVDINSFWYSFS----FARG-IDDDCFAPAQVFAVMQAWAREHGLYEHIRFNTVEVSDAHWDDESTQRWQLL	194
CtCPMO	RKLQ-YKVHLVDAAGDIDGIIWHWNCYPGARVDTHCQIYQYS-IPELWQEFNWKLEFPNWAQMRVYFHFADKLDLSKDISFNTRVQSAVDEGTREWTVR	138
PAMO	RELQ-RSVHVVIETAGDVGGMVYWRYPGARCDIESIEYCYFSSEEVLQEWNTERYASQPELRLYINFAVDKEDLRSQIETHHTVTAATAAFDEATNTWTV	134
CHMO	TTTG---QYLSKTVIFA-TGPITEAQIPRLEGLDTFKGEMFHSKWNHD-YDLTGKRIAVIGTASAIQFVPQIQPKAKELLVFQRTAPWVLPKPDVT---	228
CPMO0	MADG---KVVRRARVLVSGMGLSRSPYNIPGLDSEFEGECFHSQHWNHA-YALAGKRVAIVIGTASAIQFVPQIAIPQAQLDLYQRTPPWIMPKPDR---	225
HAPMO	YRDSGGTQVDSNVVVFVAVGQLNRPMIPAIPIGIEETFKGPMFHSAQWHD-VDSWGRKRVGVIIGTASATQFIPQLAQTAAELKFVARTTNWLLPTDLHE-	292
CtCPMO	SIGH---QPIQARFVIANLGFASPSPTPNVDGIEETFKGQWYHTALWPQEGVNMAGKRVAIIGTSSGQVVAQEAALDAKQVTVYQRTPNLALPMHQKLS	235
PAMO	TNHG---DRIRARVLIMASGQLSVPQLPNFPGKDFAGNLTGNWPHPEPVDVFSGRVGVIGTSSGSIQVSPQIAKQAELLVFQRTPHFAVPARNAPLD	231
CHMO	-----DLGEFSKSVIAKYPISQASWRKSVALTLNAINFG-LRNPLALKPVNVLGKQ-----LLKLQIN	285
CPMO0	-----EVSATERWFLRRLPFTQTLVTRTGIWYVLLSERVLGFVLPKIMKVVERR---	283
HAPMO	-----KISDSCKWLLAHVPHYSLWYVAMKQLSAGFLEDDVVDVGGTPELAVSARNDRLLRQDISAWMEPQAD	362
CtCPMO	AEDNLRMKPELPAAFERRRKGCFAGDFDFITAKNATELSAERTEILEELWNASGFRYLANFQDYVDFDDKANDVYVEFRWQ-----KVRARIK	323
PAMO	PEFLADLKKRYAEFREESRNTPGGTHRYQPKSALEVSDEELVETLERYWQEGP-DILAAVRDILRRDRANERVAEFIRN-----KIRNTRV	318
CtCPMO	-----MTTMTTTEQLGMNNSVNDKLDVLLIGAG	30
PAMO	-----MAGQTTVDSRRQPP-----EEVDVLVVGAG	25
HAPMO	MDLETEAQQVREFCLQLRIDFRDSGQPAPGRPTSDQLHILGTLWLMGVPVIEPYLPLTAAEEAVTAEEDLRAPRWHKDHVASGRDFKVVIIIGAG	90
CHMO	-----MTAQMKTNASAKKAVNSPNHIFDTFIVGAG	30
CPMO11	-----MTASSPDD-----AVLDVLVVGAG	19
CtCPMO	FTGLYQLYLRKLGK--VHLVDAGADIGGIIWHWNCYPGARVDTHCQIYQYS-IPELWQEFNWKLEFPNWAQMRVYFHFADKLDLSKDI	117
PAMO	FSGLYALYRLRELGRS--VHVIETAGDVGGVWYWRYPGARCDIESIEYCYFSSEEVLQEWNTERYASQPEILRYINFAVDKFDLRSGI	113
HAPMO	ESGMIAALRFKQAGVP--FVIEYKGNVDVGGTWTRENTYPGCRVDINSFWYSFS----FARGIWDCCFAPAQVFAVMQAVAREHGLYEH	173
CHMO	ISGIATAIRLDQVGYV--YKIIIEKASRVGGTWTRENTYPGCGDVPVLSYYS----FAPSAKWLSHLFARQPELTVLEDSVNEFNVASKI	115
CPMO11	LSGIGARQLQMRCPKRYAILEAREAMGGTDLFRYPGIRSDSOMTYLGYR----FKPWRGAKAIADGSPISIRAYIRETAAEAGITSHI	104
CtCPMO	SFNTRVQSAVDFEGTREWTVRS---IGHQPIQAR---FVIANLGFASPSPTPNVDGIEETFKGQWYHTALWPQEGVNMAGKRVAIIIGTGS	200
PAMO	TFHTTVAATAAFDEATNTWTVDT---NHGDRIAR---YLIMASGQLSVPQLPNFPGKDFAGNLYHTGNWPHPEPVDVFSGRVGVIGTGS	196
HAPMO	RFNTEVSDAHWDDESTQRWQLLYRDSGEGTQVDSN---VVVFVAVGQLNRPMIPAIPIGIEETFKGPMFHSAQWHD-VDSWGRKRVGVIIGTAS	258
CHMO	EFNNELLNAAWDSRHLWVLDT---TTGQYLSK---TVIFATGPITEAQIPRLEGLDTFKGEMFHSKWNHD-YDLTGKRIAVIGTGA	196
CPMO11	REGHRVISAAWDSAAACWTEAERSADRSLRLRLRTRLLVYVCAAGYVSAEGRPEFAGEAQFRGSMVHPQWDES-LDYAGKRVVVIGSGA	193
CtCPMO	SGVQVAQEAALDAKQVTVYQRTPNLALPMH-QKQLSAEDNLRMKPELPAAFERRRKGCFAGDFDFIAKNAATELSAERTEILEELWNASG	289
PAMO	SQIQVSPQIAKQAELVFQRTPHFAVPAR-NAPLDPEFLADLKKRYAEFREESRNTPGGTHRYQPKSALEVSDEELVETLERYWQEGG	285
HAPMO	SATQFIPQLAQTAAELKFVARTTNWLLPT-DLHEKISDSCKWLLAHVPHYSLWYRVAMAMPQSVG-----FLEDVNVVDVGYPP	336
CHMO	SAIQFVPPQIQPKAKELLVFQRTAPWVLPKP-DTDLGEGFSKSVIAKYPISQASWRKSVALTLN-----	257
CPMO11	TAVTLVPAAMAKSAAHVMTLQRSPTYIVTRPGEADIAIHKLRVRLPERLAYAATRWNKVVLLGMMFFQLAR-----	261
CtCPMO	MTTMTTTEQLGMNNSVNDKLDVLLIGAGFTGLYQLYLRKLGK--VHLVDAAGADIGGIIWHWNCYPGARVDTHCQIYQYS-IPELWQEFNWKLEFPNWA	98
PAMO	MAGQTTVDSRRQPP-----EEVDVLVVGAGFSGLYALYRLRELGRS--VHVIETAGDVGGVWYWRYPGARCDIESIEYCYFSSEEVLQEWNTERYASQ	94
HAPMO	TAEEDLRAPRWHKDHVASGRDFKVVIIIGAGESGLYIALRLQDQVGYTYKIIIEKASRVGGTWTRENTYPGCGDVPVLSYYS----FAPSAKWLSHLFARQ	96
CHMO	MTAQMKTNASAKKAVNSPNHIFDTFIVGAGTSGIATAIRLDQVGYTYKIIIEKASRVGGTWTRENTYPGCGDVPVLSYYS----FAPSAKWLSHLFARQ	96
CPMO13	-QAGFVNRQPVLQRGRPTERIELAIIIGAGMSGIAAIIQAKDRGR-FRIFDSNNKQVWMAANDYPGVAVDTPTAYYLS-----YELNPSWSNYVPVGS	94
CtCPMO	QMRVYFHFADKLDLSKDISFNTRVQSAVDFEGTREWTVRS---IGHQPIQARFVIANLGFASPSPTPNVDGIEETFKGQWYHTALWPQEGVNMAGKRVA	194
PAMO	EILRYINFAVDKFDLRSGIETHHTVTAATAAFDEATNTWTVDT---NHGDRIARYLIMASGQLSVPQLPNFPGKDFAGNLYHTGNWPHPEPVDVFSGRV	190
HAPMO	QVFAVMQAVAREHGLYEHIRFNTVEVSDAHWDDESTQRWQLLYRDSGEGTQVDSNVVVFVAVGQLNRPMIPAIPIGIEETFKGPMFHSAQWHD-VDSWGRKRV	192
CHMO	EILSLEDSVNEFNVASKIEFNNELLNAAWDSRHLWVLDT---TTGQYLSK---TVIFATGPITEAQIPRLEGLDTFKGEMFHSKWNHD-YDLTGKRIA	190
CPMO13	EYLRYLEGIVEKHNISEFIELESEILKIQWIEEDQEWELMVKKGREASRVRAVMSCLGHLNRPNYDLPQRETFKGVSIHANRWHKHD-VDLRGRV	193
CtCPMO	IIGTGSSGQVVAQEAALDAKQVTVYQRTPNLALPMHQKLSAEDNLRMKPELPAAFERRRKGCFAGDFDFIA-KNATELSAARTEILEELWNASGFRY	293
PAMO	VIGTGSSGQVVAQEAALDAKQVTVYQRTPHFAVPARNAPLDPEFLADLKKRYAEFREESRNTPGGTHRYQPKSALEVSDEELVETLERYWQEGP-DI	288
HAPMO	VIGTGASATQFIPQLAQTAAELKFVARTTNWLLPTDLHEKISDSCKWLLAHVPHYSLWYRVAMAMPQSVGF-LEDVNVVDVGYPPTELAVSARNDRLRQ	291
CHMO	VIGTGASAIQFVPPQIQPKAKELLVFQRTAPWVLPKPDTLG-EFSKSVIAKYPISQASWRKSVALTLNAINF-GLRNLPAKPVNVLGKQLLKLQIN---	285
CPMO13	IIGTGATGQVVIKISQVDHLETFEMRQPMWIIPNQAAGDGIQPSKRWARQYLPYFLHWRDLKSYWYTSOKVGYPIVRADPEWAKTHVSPANDRMLF	293
CtCPMO	LAAFQDYLFDKANDVYVEFRWDRKVRARIKDPKVAEKLAPMHPKPYGAKRPSLEQWYEIFNQNNVTLVDVNETPVLRITEKGIPTAE-GEAEFDLIV	392
PAMO	LAAVRDILRRDRANERVAEFIRNKIRNTVTRDPEVAERLVPKG--YFPGTKRLLILEIDYEMFNDRNVHLVDTLSAPITETIPRGVRTSE-REVEYDLSVL	385
HAPMO	ISAWMEPQAD-----RPDLREVLIPTD-----PVGGKRIVRDNGTWISTLKR---DNVSMIRQPEIETPKGICCVDDGETHEFDLIV	367
CHMO	-----DPVLRKNVTFNF-----TGKRILFANNYPALQSP---NTLIPHGLVKEGNTVVAANGERHEVDVII	349
CPMO13	CINYNISCFGE-----GSELARKETPYAPLGKRSVRDPYDFAPGGFYVALSQPNVALETSKILARVVPDGLITVDGKLIELDVYII	374

Figure 54. Multiple nucleotide sequence alignment of CPMO0, CPMO11, and CPMO13 from *C.necator* H16 with known BVMOs.

CHMO: cyclohexanone monooxygenase from *Acinetobacter calcoaceticus*.^[203] HAPMO: 4-hydroxy acetophenone monooxygenase from *Pseudomonas fluorescens*.^[204]

CtCPMO: cyclopentanone monooxygenase from *Comamonas testosteroni*.^[205] PAMO: phenylacetone monooxygenase from *Thermobifida fusca*.^[206] Alignment created with CLUSTALW. The identical residues are aligned manually and visualized with SnapGene software.^[207] Similar residues are highlighted in grey. The rest alignments of BVMOs in this study were presented in Figure 64-Figure 80.

As shown in Table 29, the sequence identity with CHMO was determined to be 51% for CPMO0, 17% for CPMO11, and 20% for CPMO13. The sequence identity with HAPMO was determined to be 27% for CPMO0, 15% for CPMO11, and 32% for HAPMO. The sequence identity with CtCPMO was determined to be 18% for CPMO0, 14% for CPMO11, and 20% for CPMO13. The sequence identity with PAMO was 25% for CPMO0, 16% for CPMO11, and 28% for CPMO13. CPMO0 and CPMO13 presented higher sequence similarity with BVMOs comparing with CPMO11. From the alignment presented in Figure 54, the sequence motifs (FxGxxxHxxxW) indicate catalytic activity, and the motifs for FAD coenzyme binding (two GxGxxG/A) were highly conserved in these three putative BVMOs. Further studies were focused on CPMO0 firstly.

Table 29. Pairwise sequence identities between CPMO0 and CPMO11 from *C.necator* H16 with typical BVMOs.

Enzyme	CHMO	HAPMO	CtCPMO	PAMO
CPMO0	262/507* 51%	157/580* 27%	100/550* 18%	136/542* 25%
CPMO11	89/507* 17%	89/580* 15%	79/550* 14%	89/542* 16%
CPMO13	103/507* 20%	186/580* 32%	110/550* 20%	151/542* 28%





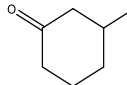
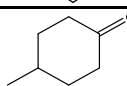
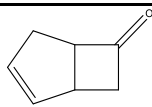
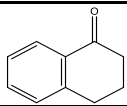
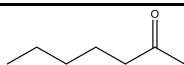
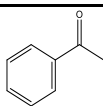
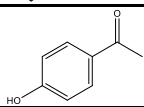
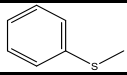
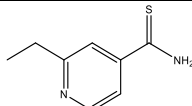
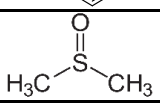
*Number of identical amino acids/Number of aligned amino acids

3.2. Determination of substrate scope for CPMO0

To probe CPMO0 can catalyze Baeyer-Villiger oxidations, the CPMO0 overexpressed whole-cell was used as a biocatalyst to perform the initial substrates screening of CPMO0. Fourteen known BVMO substrates were tested at a concentration of 10 mM. As shown in Table 30, the substrates are Alicyclic ketones (Cyclobutanone (1), Cyclopentanone (2), Cyclohexanone (3)). Substituted alicyclic ketones (2-methyl Cyclohexanone (4), 3-methyl Cyclohexanone (5), 4-methyl Cyclohexanone (6)). Bicyclic ketones (Bicyclo[3.2.0]hept-2-en-6-one (7), Alpha tetralone (8)). Aliphatic ketones (2-heptanone (9), Acetophenone (10), 4'-hydroxy Acetophenone (11)). Sulfides (Thioanisole (12), Ethionamide (13), Dimethylsulfoxide (14)).

The reactions were performed with 10 mg wet whole cells under oxygen flow for 48 h at 25 °C. The products were determined via GC-MS analysis. As shown in Table 30, among the substrates, three types of substrate were converted into their corresponding products, which are Cyclopentanone of Alicyclic ketones, Bicyclo[3.2.0]hept-2-en-6-one of Bicyclic ketones, Thioanisole and Dimethylsulfoxide of Sulfides. No products for substituted alicyclic ketones and aliphatic ketones were detected.

Table 30. Substrates scope screening of CPMO0.

Number	Substrates	Activity
Alicyclic ketones		
1	Cyclobutanone 	0
2	Cyclopentanone 	Have activity
3	Cyclohexanone 	0
Substituted alicyclic ketones		
4	2-methyl Cyclohexanone 	0
5	3-methyl cyclohexanone 	0
6	4-methyl cyclohexanone 	0
Bicyclic ketones		
7	Bicyclo[3.2.0]hept-2-en-6-one 	Have activity
8	Alpha tetralone 	0
Aliphatic ketones		
9	2-heptanone 	0
10	Acetophenone 	0
11	4'-hydroxyacetophenone 	0
Sulfides		
12	Thioanisole 	Have activity
13	Ethionamide 	0
14	Dimethylsulfoxide 	Have activity

3.3. *In vivo* conversion of Cyclopentanone to δ -valerolactone by CPMO0

Previously it was suggested that limited stability compromised the performance of BVMOs *in vitro* activity. The higher stability may be benefited when performing oxidation reaction *in vivo* since the efficient NADH regeneration dehydrogenases from *C.necator* H16 enable to recycle of NADH, which is the prerequisite for BVMOs. Therefore, *C.necator* H16 whole-cell-mediated biotransformations were performed.

The CPMO0 overexpressed and the green fluorescent protein (GFP) overexpressed *C.necator* H16 whole cell were used as the catalyst. The reactions were performed with 10 mg wet whole cell under oxygen flow for 48 h at 25 °C. As shown in Figure 55, the CPMO0 overexpressed *C.necator* H16 whole cell converted 10 mM cyclopentanone to 0.93 mM δ -valerolactone, which was nine-fold higher compared to 0.13 mM δ -valerolactone produced by the control cells harboring the green fluorescent protein (GFP). This demonstrated that CPMO0 overexpressed *C.necator* H16 strain would be used as the cell factory to produce high-value chemicals.

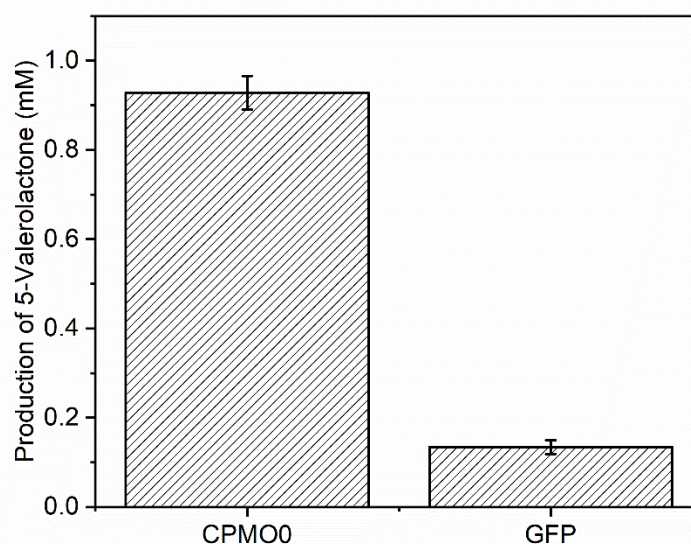


Figure 55. The whole cell-mediated biotransformation from Cyclopentanone to 5-Valerolactone.

Reaction conditions: Whole cell=10 mg; Reaction buffer: Potassium phosphate buffer (50 mM, pH 8); Reaction volume=1000 μ L; Substrates=10 mM. 25 °C; reaction time: 48 h.

3.4. Heterologous expression, purification, and activity study of CPMO0 in *E.coli*.

E.coli is the dominant host for producing recombinant proteins. It presented various advantages comparing with other microbial platforms: the genetics of *E.coli* are well-characterized, growing fast and inexpensive, high yield of protein production, and variety

of available molecular tools for *E.coli*.^{[208],[209]} Therefore, we attempted to express and purification of CPMO0 in *E.coli* for further study.

The gene CPMO0 was amplified from *C.necator* H16 by PCR using the isolated genomic DNA as a template and inserted into the IPTG induce vector pET28a(+). The construct pET28a-His-CPMO was designed with a His-tag fused to the *N*-terminus of CPMO0. Protein expression was performed in *E. coli* BL21 (DE3). Different expression temperatures were investigated, and the amount of CPMO0 in the soluble and inclusion body fraction was studied by SDS-PAGE.

As presented in exemplary SDS-PAGE gel (Figure 56), the protein CPMO0 was observed in the insoluble fractions with masses of approximately 60 kDa. When decreased expression temperature from 28 °C to 18 °C (IPTG 0.1 mM), the protein still was observed in the insoluble fractions. We attempted to purify the CPMO0 protein from soluble fractions with Ni-affinity chromatography. Unfortunately, no soluble form of CPMO0 under the assay conditions was detected.

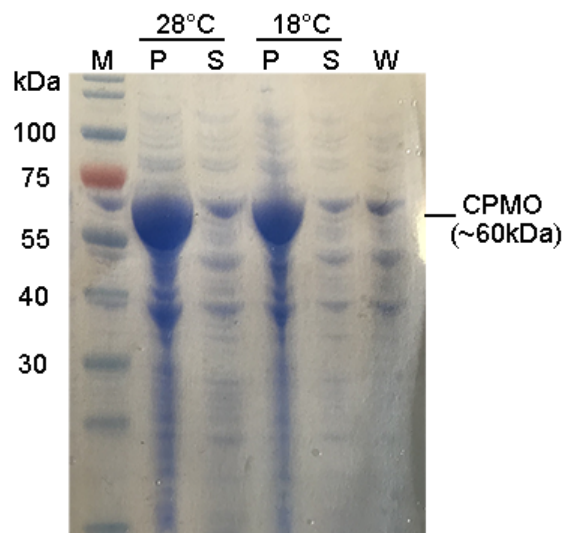


Figure 56. Expression analysis of CPMO0.

Line 28 °C p-s: after inducing, the protein was express at 28 °C, insoluble (p) and soluble (s) fraction of CPMO. Line 18 °C p-s: after inducing, the protein was express at 18 °C, insoluble (p) and soluble (s) fraction of CPMO. W: Whole cell without inducing. M: Protein ladder.

We attempted to obtain the recombinant His-tagged CPMO0 from the inclusion body. The unfolding and refolding steps are described experimental section. Results are presented in Figure 57. We could observe for line 3 and line 4 after the unfolding step followed with the centrifuge, and almost all protein remained in the insoluble fraction. No target bands from the soluble fraction were observed. We continued refolding and dialysis process overnight. Numerous bands in the solution were observed in line 2. We used this solution as the catalyst to perform the oxidation reaction. There were no products detected. We envisioned that either the efficiency of refolding was low or the process inactivated the

protein. So the other solutions to obtain a functional soluble CPMO0 enzyme were attempted.

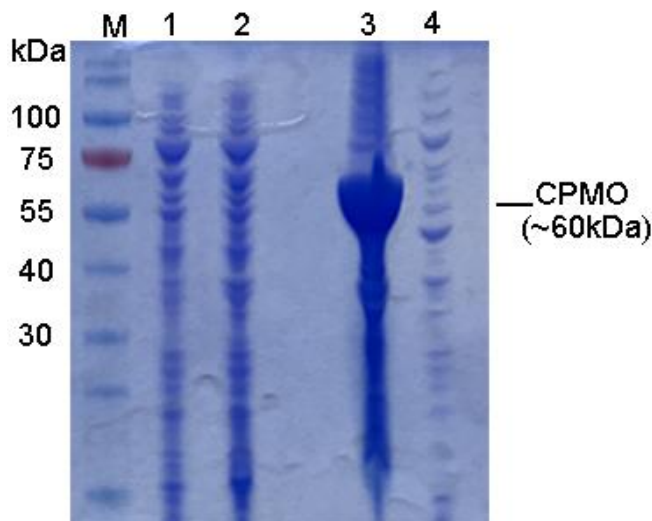


Figure 57. SDS-PAGE analysis of CPMO0 solubilization.

Line 1: concentrated fractions after refolding CPMO; Line 2: concentrated fractions after dialysis overnight; line 3: the insoluble fractions after unfolding. Line 4: the soluble fractions after unfolding.

It has already been reported that co-expression of molecular chaperones in the *E.coli* cytoplasm could significantly increase the yield of soluble protein.^{[210]–[212]} It was studied clear recently, in *E.coli* cells, the *in vivo* protein folding is an energy-related process mediated by two types of folding modulators. DnaK-DnaJ-GrpE and GroES-GroEL systems are the molecular chaperones that bind to partially folded proteins, inhibit aggregation reactions, and facilitate folding properly by their catalytic functions.^[213]

Then, we attempted to co-express CPMO0 with a chaperon GroES-GroEL system to further improve CPMO0 solubility. The *E.coli* DH5 α strain harboring plasmid pGro7 was transformed with the pET28a-His-CPMO0 plasmid. The expression was analyzed by SDS-PAGE (Figure 58). The target band could be observed in the soluble fraction part. Due to the protein molecular weight of the chaperon enzyme (60 kDa) was close to the CPMO0 (60 kDa), we attempted to purify the CPMO0 protein with nickel affinity chromatography. Unfortunately, we could not obtain the purified soluble CPMO0.

We used soluble fractions as the catalyst to perform the oxidation reaction. No products could be detected. We envisioned co-expression chaperon GroES-GroEL did not improve CPMO0 solubility, or the catalytic properties of soluble recombinant CPMO0 were inhibited. A study demonstrating the co-expression of folding modulators might improve solubility, but catalytic properties were not the same as the natural enzyme.^[212]

Table 31. Chaperone Plasmid pGro7

Plasmid	Chaperone	Promoter	Inducer	Resistant Marker
pGro7	GroES-GroEL	Arab	L-Arabinose	Cm

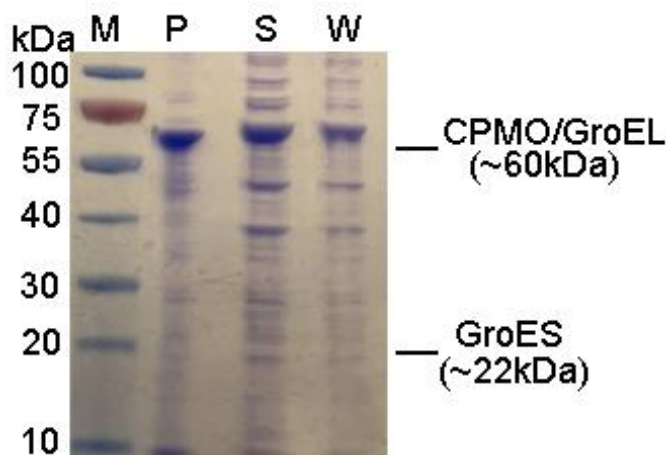


Figure 58. SDS-PAGE analysis co-expression CPMO0 with chaperon protein in *E.coli*. Line M-W: insoluble (P), soluble (S) fraction fractions, and whole cell (W) of CPMO0 with [pGro7].

As we presented previously, with recombinant plasmid pET28a-His-CPMO0, the His-tag was tightly fused to CPMO0 at *N*-terminal. The recombinant region could not code soluble protein properly. Since the fused tags might promote different effects when located at the *N*- or *C*-terminus of the passenger protein.^{[214], [215]} Therefore, for redesigned construction, the His-tag was fused at the *C*-terminal of CPMO0 was considered. On the other hand, few additional amino acids between the His-tag and recognition site of CPMO0 were expected to improve protein solubility. As a result of this, two constructions pET28a-His-CPMO0 and pET28a-CPMO0-His were redesigned. For construction pET28a-His-CPMO0, the CPMO0 gene was fused in the frame to six histidine codons already exist in the vector. For construction pET28a-CPMO0-His, the CPMO0 gene without its stop codon was fused at the *N*-terminal of six his codons already exist in the vector.

The redesigned constructions were transformed and expressed in *E.coli* BL21(DE3) with the same conditions. Different expression temperatures were investigated, and the amount of CPMO0 in the soluble and inclusion body fraction was studied by SDS-PAGE. (Figure 59) For expression at 28 °C, the recombinant enzymes were expressed successfully, but the enzyme was deposited in the inclusion body again. Decreasing the expression temperature to 18 °C could not increase CPMO0 solubility. All the soluble fractions were purified with Ni affinity chromatography, with no observable soluble target bands from SDS-PAGE.

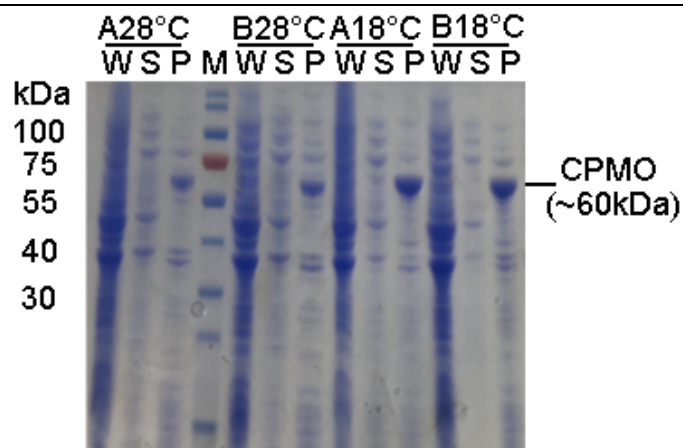


Figure 59. SDS-PAGE analysis of the expression of CPMO0.

Line A 28 °C W-P: *E.coli* harboring expression vector pET28a-His-CPMO after inducing the protein was express at 28 °C, insoluble (P), soluble (S) fraction, and whole cell (W).

Line B 28 °C W-P: *E.coli* harboring expression vector pET28a-CPMO-His, after inducing the protein was express at 28 °C, insoluble (P), soluble (S) fraction and whole cell (W).

Line A 18 °C W-P: *E.coli* harboring expression vector pET28a-His-CPMO after inducing the protein was express at 18 °C, insoluble (P), soluble (S) fraction, and whole cell (W).

Line B 18 °C W-P: *E.coli* harboring expression vector pET28a-CPMO-His, after inducing the protein was express at 18 °C, insoluble (P), soluble (S) fraction and whole cell (W).

In recent years, multiple fusion tags have been developed for recombinant protein production, such as affinity tags for protein detection and purification, enhancing protein solubility, tag removal, or applications.^{[216]–[221]} Fusion protein or peptide is a common strategy to overcome expression obstacles. In this method, the enzyme of interest but which is difficult to produce is fused to one or more other proteins to drive the resulting expression. Few kinds of fusion proteins have been used for enhancing protein solubility, such as glutathione S-transferase (GST), maltose-binding protein (MBP), thioredoxin A (TrxA), small ubiquitin-related modifier (SUMO) et al.^[222]

In our study, we selected thioredoxin A to improve the solubility of CPMO0 in *E.coli*. As one of the thioredoxins in *E.coli*, TrxA is an 11.6 kDa protein that presented high solubility in the *E.coli* cytoplasm. TrxA has been used as an N- or C-terminal fusion protein to increase soluble protein expression of proteins.^{[223]–[225]} The plasmid pET28a-His-TrxA-CPMO0 was constructed and transformed with *E.coli* BL21(DE3). After cultivation and induction, the protein was presented at soluble fractions and purified with Ni affinity chromatography (Figure 8).

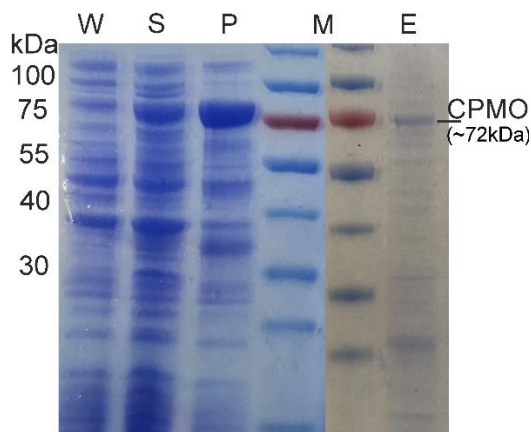


Figure 60. SDS-PAGE analysis of the expression of CPMO0 with soluble tag: TrxA in *E.coli*.

Line W-E: whole cell without inducing (W), the soluble fraction (S) after induce, insoluble fractions after induce (P), purified His-TrxA-CPMO0 enzyme (E).

The next step was activity study. The whole-cell, cell extract, and purified enzyme were used as catalysts for biotransformation of different substrates. However, no products were observed. Following optimizations of reaction conditions were performed: increasing the concentration of catalyst, adding oxygen flow, and adding phosphite dehydrogenase (PTDH) for NADPH regeneration. Unfortunately, there was no enzymatic activity could be detected from all reactions.

The loss of the cofactor caused the BVMOs loss of activity during purification, and this phenomenon had been reported.^{[226],[227]} We envisioned addition of FAD after enzyme purification might recover the activity of CPMO0. However, incubating with FAD did not result in activity recovery. Another possibility for producing inactive enzymes might be that *E.coli* produces an insufficient amount of FAD during the cultivation period. The enzyme

is inactive because of the lack of the cofactor.^{[228],[229]} From our results, the addition of FAD into culture during cultivation did not increase the enzymatic activity. There was still no production of products that could be observed to perform the reaction with the whole cell and cell free extract as catalysts. Therefore, the fusion tag has improved the solubility of CPMO0, but the soluble enzyme was inactive form.

E.coli is the dominant host for producing recombinant proteins. It presented various advantages comparing with other microbial platforms. However, the lack of sophisticated machinery to perform posttranslational modification is the main drawback for *E.coli* to expressing more complex proteins. This limitation always results in poor solubility of the target protein, which is produced as inclusion bodies.^[209] On the other hand, the strong T7 promoter drives a high protein synthesis rate, which rendered the nascent CPMO0 polypeptide trapped in an incorrectly folded state, leading to the formation of inclusion bodies.^[230] In our case, we could not to obtain the active soluble form of CPMO0 in *E.coli* with few approaches.

3.5. Overexpression and purification of CPMO0 in *C.necator* H16.

Therefore, overexpression and purification of CPMO0 in *C.necator* H16 were attempted. As Gruber *et al.* study, they designed an inducible expression vector for improved protein production in *C.necator* H16. The vector was designed based on the *j5* promoter and the *E.coli* *lacI* regulatory elements.^[200] The recombinant plasmid pKRL-CPMO0-His was successfully generated using the methodology described in experimental section 5.6. The obtained construct was confirmed by colony PCR. The confirmed construct was conjugated into *C.necator* H16 strain as described in the experimental section 5.11. The success of conjugation was confirmed by isolating plasmid as the template for colony PCR.

As shown in Figure 61, the expression of CPMO0 in *C.necator* H16 resulted in 7.8 mg of purified (Ni-affinity chromatography) protein per 500 mL cultivation.

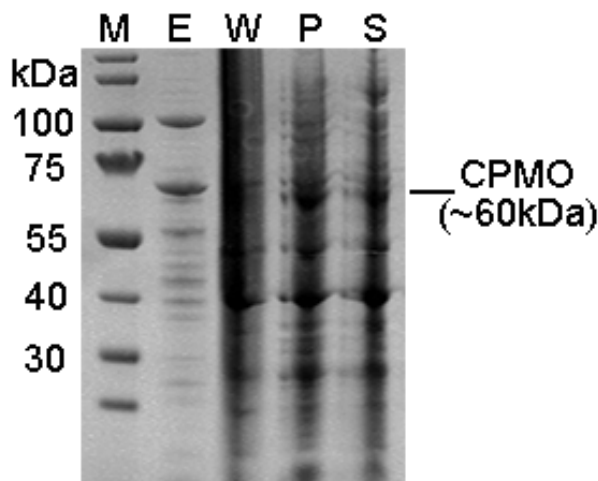


Figure 61. SDS-PAGE analysis of the expression of CPMO0 in *C.necator* H16. Line E-S: *C.necator* H16 harboring expression vector pKRL-CPMO0-His, after inducing with 0.1 mM IPTG, the protein was expressed at 25 °C, purified CPMO0 enzyme (E), whole cell (W), insoluble (P), and soluble (S) of CPMO0.

3.6. *In vitro* activity assessment of purified enzyme of CPMO0

To validate purified CPMO0 is the functional enzyme, screening for the substrate preference profile of CPMO0 was performed with purified enzyme as the catalyst. The same thirteen compounds were selected, as shown in Table 30. The products were determined via GC-MS. Reactions were performed in 2 ml glass vials containing 10 μ M enzyme, 10 mM substrates, and 10 mM NADPH at 25 °C with horizontal shaking to increase oxygen insertion into the solution. Reaction times were set to 1 h.

Till now, Bicyclo[3.2.0]hept-2-en-6-one, Thioanisole, and Dimethylsulfoxide were converted into corresponding products when the purified enzyme was the catalyst. There was no detectable activity for Cyclohexanone, which was concerted when using the whole-cell as the catalyst. On the one hand, the whole cell-mediated biotransformation is influenced by the amount of functional enzymes, cell wall transport, substrate acceptance, or other factors. On the other hand, enzyme activity is measured *in vitro* under the same conditions that consistently do not resemble *in vivo*.^[231] Thus, further characterization of CPMO0 would favor the understanding of this novel enzyme.

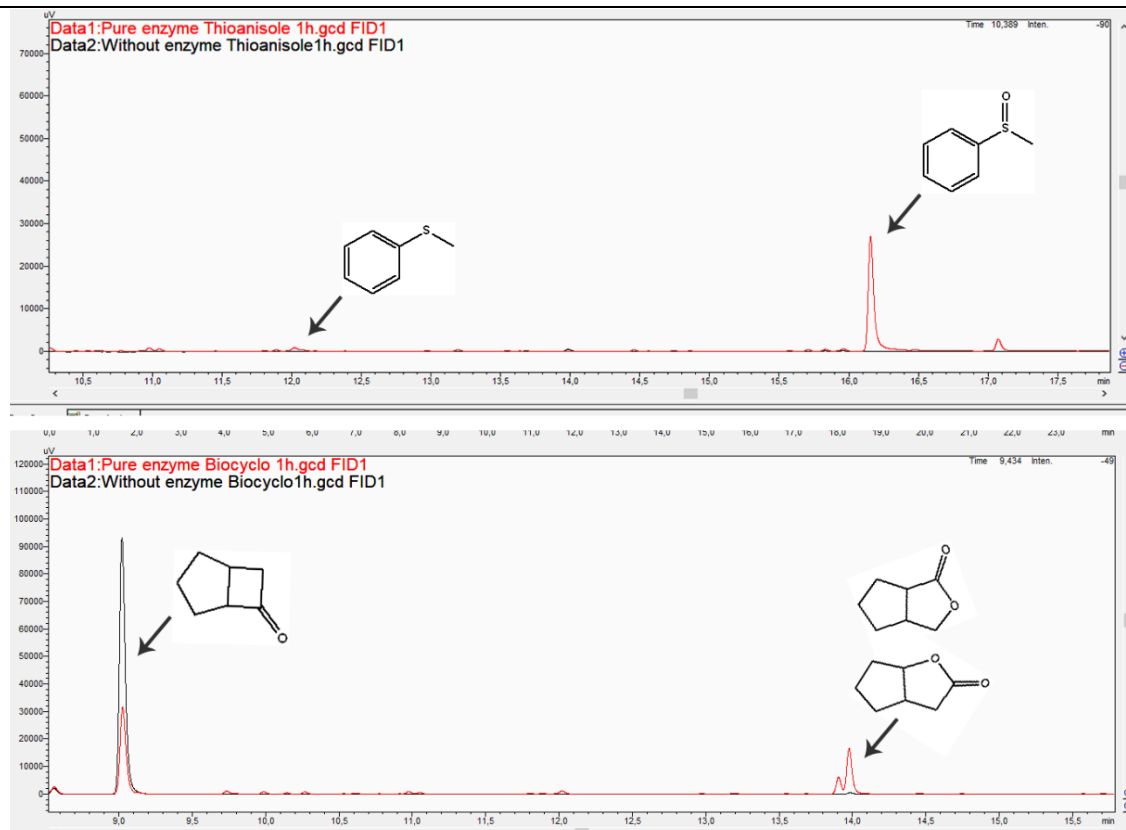


Figure 62. GC-FID Chromatogram of substrates and products.

3.7. Melting Temperature determination of CPMO0 by ThermoFAD

Enzymes perform functions in specific cellular environments. The biochemical characteristics are essential when the purified enzyme was utilized as a catalyst for biotransformation *in vitro*.^[232] Stability is an important characteristic related to the temperature-dependent properties of enzymes. Validating the stability parameters of enzymes may give more insight into the suitability for *in vitro* biotransformation.^[233]

The thermostability of CPMO0 in Potassium phosphate buffer with different pH values and the effect of the addition of cosolvent was assessed via the *ThermoFAD* method.^[110] This method is based on the fact that the flavin-containing proteins display different fluorescent properties between the folded and denatured state. In particular, during the denaturation of flavoproteins, the secondary and tertiary structure of the protein is disrupted, and interactions with the flavin break down. The released free flavin usually results in a significant increase in fluorescence at 530 nm.^[110]

The comparison of melting temperatures (T_m) for CPMO0 was presented in Figure 63. In potassium phosphate buffer, the T_m of pH 10 and pH 8 was ~ 10 °C higher than pH 6. Meanwhile, CPMO0 melting temperature above 78 °C indicated that CPMO0 might be a relatively stable enzyme in potassium phosphate buffer.

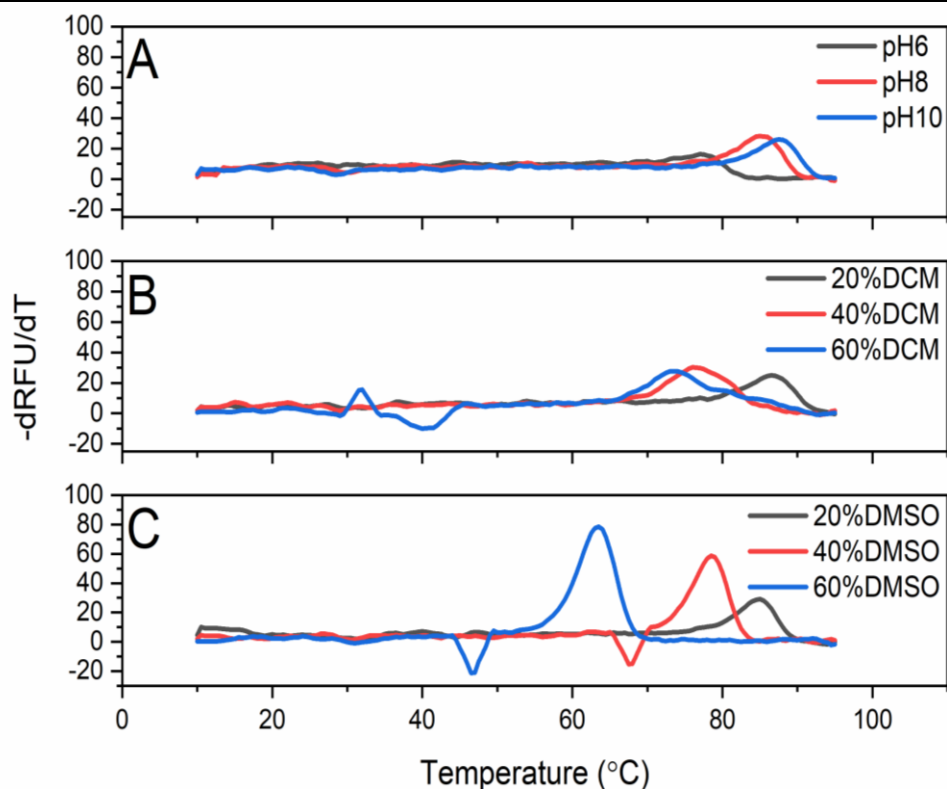


Figure 63. The first derivative curves of the melting temperature curves were obtained in fluorescence measurements.

When organic cosolvent into the reaction system, both 20% of DCM and DMSO exhibited no influence for CPMO0 thermostability. T_m of 60% DCM decreased 10 °C compare with 20% DCM. When up to 60% DMSO, the T_m was sharply decreased 22 °C comparing with 20% DMSO. Since BVMOs enable oxidize DMSO, the reason for T_m significant decreasing when DMSO presented in the system might be caused by DMSO reaching the CPMO0 substrate activate the site.

Table 32. Comparison of T_m determined by ThermoFAD for CPMO0.

	$T_m(^{\circ}\text{C})$
Potassium phosphate Buffer (100 mM) pH 6	78
Potassium phosphate Buffer (100 mM) pH 8	85
Potassium phosphate Buffer (100 mM) pH 10	87
Potassium phosphate Buffer (100 Mm; pH 8)+20% DCM	85
Potassium phosphate Buffer (100 Mm; pH 8)+40% DCM	78
Potassium phosphate Buffer (100 Mm; pH 8)+60% DCM	75
Potassium phosphate Buffer (100 Mm; pH 8)+20% DMSO	85
Potassium phosphate Buffer (100 Mm; pH 8)+40% DMSO	79
Potassium phosphate Buffer (100 Mm; pH 8)+60% DMSO	63

After this assessment, it was suggested that CPMO0 was thermostable when incubated in the phosphate buffer with pH 8. The cosolvent used for dissolving substrates, CPMO0 is more thermostable in DCM than DMSO.

4. Conclusion and outlook

In this study, we analyzed eighteen putative BVMO encoding genes from organism *C.necator* H16. Three of them were identified as Type 1 Baeyer-Villiger monooxygenase encoding genes. One of these three putative BVMOs, CPMO0, was selected for further studies.

After screening five types of BVMO's substrates, three types of substrates were converted into their corresponding products, which are Cyclopentanone of Alicyclic ketones, Bicyclo[3.2.0]hept-2-en-6-one of Bicyclic ketones and Thioanisole, Dimethylsulfoxide of Sulfides. This confirmed that CPMO0 is a BVMO.

From the results obtained till now, the substrate specificity of CPMO0 provides some interesting results. For alicyclic ketone substrates, CPMO0 only converted Cyclopentanone but no activity for Cyclohexanone at all. It was interesting because CPMO0 showed 51% amino acid identity with CHMO (cyclohexanone monooxygenase) from *A.calcoaceticus*. For bicyclic ketone substrates, CPMO0 showed activity for Bicyclo[3.2.0]hept-2-en-6-one, which is the model substrate to study regio/stereoselectivity of BVMOs.

After performing many approaches to heterologous expressed CPMO0 in *E.coli*, no active and soluble protein could be obtained. Therefore, we attempted to express CPMO0 in *C.necator* H16, which lead to obtaining functional and soluble CPMO0 eventually. The purification yield was relatively low. The induction with 0.1 mM IPTG resulted in 7.8 mg of purified (Ni-affinity chromatography) protein per 500 mL cultivation. Thus, the expression conditions could still be optimized for improving expression levels.

For the outlook of this project:

Part 1. Baeyer-Villiger monooxygenases enable to catalyze the reactions that are hard to achieve through chemical methods. The CPMO0 was experientially confirmed as BVMO. The other two putative proteins CPMO11 and CPMO13, which have a higher potential are Type 1 BVMO as well. The rest sequences also contain BVMO identical motifs partially. The set of studies of these sequences from *C.necator* H16 would get insight into the structure or function of BVMOs.

Part 2. As a novel enzyme, the catalytic activity and selectivity of CPMO0 still present enormous potential to explore. Meanwhile, the biochemical characterization of the CPMO0, such as optima of pH, temperature, and solvent tolerance *et al.*, will be needed.

Part 3. Our results suggested that biocatalytic systems employing *C.necator* H16 whole cells can be used to circumvent expensive coenzyme regeneration procedures of CPMO0. Moreover, *C.necator* H16 can grow to high-cell densities. Therefore, *C.necator* H16 was a suitable microbial platform to produce BVMOs and the high-value products oxidized by BVMOs on a large scale, which would favor industrial applications of BVMOs.

5. Experimental section

5.1. Bacterial strains, plasmids, and primers

All strains and plasmids used in this study were listed in Tables 10 and 12.

Table 33. Strains used in this study

Strains	Purpose	Reference and source
<i>E. coli</i> TOP10	Cloning experiments	Thermo Scientific
<i>E. coli</i> BL21 (DE3)	Production of recombinant proteins	Thermo Scientific
<i>E. coli</i> S17-1	Used as the donor strain in conjugation experiments	Invitrogen
<i>C.necator</i> H16	Wildtype strain	DSMZ 428

Table 34. Plasmids used in this study

Plasmid	Description	Reference and source
pKRL-P _{j5} -GFP	Positive control for lac inducible system with the gene encoding for CPMO expression	[200]
pKRL-P _{j5} -CPMO0	Lac inducible expression system with gene encoding for CPMO expression	This study
pKRL-P _{j5} -CPMO0-His	Lac inducible expression system with gene encoding for CPMO expression	This study
pET28a-His-CPMO0	Lac inducible expression system with gene encoding for CPMO expression	This study
pET28a-His-CPMO0 (natural His tag)	Lac inducible expression system with gene encoding for CPMO expression	This study
pET28a-CPMO0-His (natural His tag)	Lac inducible expression system with gene encoding for CPMO expression	This study
pET28a-TrxA-His-CPMO0 (natural His tag)	Lac inducible expression system with gene encoding for CPMO expression	This study
pGro7	L-Arabinose inducible expression for chaperone molecular GroES-GroEL	Thermo Scientific

Table 35. List of primers used in this study.

Name	Sequence (5'-3')
pKRL-P _{j5} -CPMO0_fdw	ATGAATGCACCCGCGGAGCTGC
pKRL-P _{j5} -CPMO0_rev	TCAGGCACCGGCGACCGTG
pKRL-P _{j5} -CPMO0-His insert_fdw	CTTTAAGAAGGAGATATACATATGAATGCAC CCGCGGAGCTGCCGC
pKRL-P _{j5} -CPMO0-His insert_rev	CCAAAACAGCCAAGCTTCAATGGTGTGGT GATGATGGGCACCGGCGACCGTGGCG
pKRL-P _{j5} -CPMO0-His vector_fdw	GTGCCCATCATCACCATCACCATTGAAGCT TGGCTGTTTTGGCGGATGAGAGAAG
pKRL-P _{j5} -CPMO0-His vector_rev	GCTCCGCGGGTGCATTCATATGTATATCTC CTTCTTAAAGTTAAACAAAATTATTTTC
pET28a-His-CPMO0_fdw	CACCATCACCATTGACGCGGATCCGCGACC CATTG
pET28a-His-CPMO0_rev	GCGGGTGCATTCATCCCAAGCTTGCGGCC GCACTCGA
pET28a-His-CPMO0 insert part (natural His tag)_fdw	GCGGATCCATGAATGCACCCGCGGAG
pET28a-His-CPMO0 insert part (natural His tag)_rev	GCTCGAGTCAGGCACCGGCGACC
pET28a-His-CPMO0 vector part (natural His tag)_fdw	GCATTCATGGATCCGCGACCCATTG
pET28a-His-CPMO0 vector part (natural His tag)_rev	GTGCCTGACTCGAGCACCACCACC
pET28a-CPMO0-His insert part (natural His tag)_fdw	CCATGGGCATGAATGCACCCGCGGAG
pET28a-CPMO0-His insert part (natural His tag)_rev	GCTCGAGGGCACCGGCGACCG
pET28a-CPMO0-His vector part (natural His tag)_fdw	GCATTCATGCCCATGGTATATCTCCTTCTTA AAGTTAAACA
pET28a-CPMO0-His vector part (natural His tag)_rev	CCGGTGCCCTCGAGCACCACCACCACCAC CACT
pET28a-TrxA-His-CPMO0 insert part (natural His tag)_fdw	GCTCCGCGGGTGCATTGGATTGGAAGTACA GG
pET28a-TrxA-His-CPMO0 insert part (natural His tag)_rev	AAGAAGGAGATATACCATGGGCATGCATCA TCACCAT
pET28a-TrxA-His-CPMO0 vector part (natural His tag)_fdw	GCTTCCGAAAATACGATTACATTCTGG
pET28a-TrxA-His-CPMO0 vector part (natural His tag)_rev	CCTGTACTIONCCAATCCAATGCACCCGCGGA GC

5.2. Gene analysis

The genomic sequences were obtained from National Center for Biotechnology Information (NCBI; <http://www.ncbi.nlm.nih.gov/gorf/gorf.html>). Pattern Hit Initiated BLAST searches were performed using the BLAST function at the NCBI website (<http://www.ncbi.nlm.nih.gov/BLAST/>). The nucleotide encoding regions in the *C.necator* H16 genome comprise the BVMOs fingerprint motifs were selected (Table 42, Table 43, and Table 44). The multiple nucleotide sequences alignment with putative BVMOS of *C.necator* H16 and known BVMOs using CLUSTALW (<https://www.genome.jp/tools-bin/clustalw>). The identical residues were aligned manually with SnapGene.

5.3. Cultivations of *E.coli* and *C.necator* H16 strains

E.coli cell was cultivated at 37 °C on lysogeny broth (LB) media with kanamycin [40 µg/ml]. *C.necator* H16 cell were cultivated at 28 °C on nutrient broth (NB) or tryptic soy broth (TBS) media supplemented with gentamicin [20 µg/ml], kanamycin[200 µg/ml] and 0.6% or 2% fructose according to application. All basic media components were purchased from Sigma-Aldrich (St. Louis, MO, USA), Carl Roth (Arlesheim, Germany) and Becton Dickinson and Company (Franklin Lakes, NJ, USA).

5.4. Glycerol stocks

Cultures were maintained in 25% (v/v) glycerol stocks at -20 °C and -80 °C. For this purpose, 50% (v/v) sterile glycerol was thoroughly mixed with the respective overnight culture in a 1:1 ratio in cryogenic tubes. Afterward, glycerol stocks were frozen and replaced after repeated use.

5.5. Isolation of *C.necator* H16 genomic DNA

To isolate genomic DNA (gDNA), overnight cultures of the respective *C.necator* H16 strain were prepared according to section 4.3 and grown to saturation. Cultures (5 mL) were harvested by centrifugation (5 000 rpm, 5 min, 4 °C). The obtained cell pellet was resuspended in TNE buffer (1 mL), and the supernatant was discarded after centrifugation (5 000 rpm, 5 min, 4 °C) before the cells were resuspended again in TNEX buffer (540 µL). Lysozyme (60 µL, 5 mg mL⁻¹ in ddH₂O) was added and the cell suspension incubated at 37 °C for 20 min. Subsequently, Proteinase K (30 µL, 20 mg mL⁻¹, Thermo Scientific) was added, and the mixture was incubated at 65 °C for 60 min. The cleared solution was extracted with PCI (600 µL; phenol/chloroform/isoamyl alcohol = 25:24:1), and phase separation was achieved by centrifugation (13 000 rpm, 10 min, 4 °C). The upper, aqueous phase (400 µL) was transferred to a fresh tube before 5 M NaCl (40 µL), and ice-cold absolute EtOH (1 mL) was added. After careful mixing, a cloudy gDNA precipitate formation was observed, pelleted by centrifugation (13 000 rpm, 20 min, 4 °C). Optionally, precipitation of gDNA could be extended overnight at -20 °C. After centrifugation, the supernatant was removed, and the DNA pellet was washed by addition of 70% EtOH (700 µL) without mixing and subsequent centrifugation (13 000 rpm, 5 min, 4 °C). The washing

step was repeated, and the resulting DNA pellet was dried at 60 °C. Genomic DNA was dissolved in ddH₂O (100 µL), RNase A (2 µL, 10 mg mL⁻¹, Thermo Scientific) was added, and the mixture was incubated for 30 min at 37 °C. The resulting DNA preparation was analyzed by agarose gel electrophoresis (3 µL) to estimate the concentration and stored at 4 °C (short-term) or -20 °C.

5.6. Plasmid construction

The plasmid pKRL-P_{j5}-CPMO0 was constructed based on the pKRL-P_{j5}-GFP backbone. pKRL-P_{j5}-CPMO0_fdw and pKRL-P_{j5}-CPMO0_rev were used to amplify the gene CPMO0 from the genomic DNA of *C.necator* H16. The PCR products and pKRL-P_{j5}-GFP were restricted with NdeI/ HindIII and combined by ligation.

The series of plasmids of pET28 were constructed based on the pET28 backbone, and the primers were presented as in Table 35, respectively. The PCR products and vector backbone was ligated with fast cloning.

The plasmids of pKRL-P_{j5}-CPMO0-His were constructed based on the pKRL-P_{j5}-GFP backbone, and the primers were presented as in Table 35, respectively. The PCR products and vector backbone was ligated with Gibson Assembly.

5.7. Fast Cloning

Prepare a map of the desired plasmid (gene of interest at the desired position of the vector of interest). Two pairs of primers need to design. Take about 15 overlap bases of the gene of interest and 15 of the vector of interest for each primer. The primers should end with G or C, especially if there are many T and A near the ends. A mix containing all reaction components was prepared as in **Table 36**.

Table 36. Components of the FastCloning PCR reactions.

2 µL	10X Phusion Buffer	95 °C	30/60 sec	Initial denaturation
0.5 µL	Primer fwd [10 µM]	95 °C	15 sec	Denaturation
0.5 µL	Primer rev [10 µM]	Var.	15 sec	
0.5 µL	DNA Polymerase (Phusion)	72 °C	var.	Elongation
2 µL	dNTPs (2mM)	72 °C	5 min	
1 µL	Plasmid DNA	4 °C	∞	Final elongation
13.5 µL	ddH ₂ O			

After completing the PCR runs, the samples are mixed with the appropriate dye, and 5 µL of each sample is loaded onto a 1% agarose gel. This is run at 120 V for a period of 60 minutes to confirm the fragments were correctly amplified. Once success amplification is confirmed, the inserts and vector need to be digested with the *DpnI* enzyme. This serves to remove methylated template DNA. 1 µL of *DpnI* is added to 20 µL each PCR amplified fragment (the vector and inserts). The vector and respective insert are then combined and placed at 37 °C for a period of 60 minutes. Following the *DpnI* digest procedure, 3 µL of

the digested sample is transformed into 50 μL of chemo competent *E. coli* cells. The transformed cells are then plated on LB-agar plates that are supplemented with kanamycin ($40 \mu\text{g mL}^{-1}$).

5.8. Gibson Assembly

Gibson assembly primers are composed of two parts: primer sequence and overlap sequence. The overlap sequence needs to have between 20 – 150 bp homology to insert or vector. If amplifying vector, digest with DpnI before Gibson assembly to remove parental vector. Cleanup is not required.

Thaw a 15 μL assembly mixture (Table 37) aliquot and keep on ice until ready to be used. Add 5 μL of DNA to be assembled to the master mixture. The DNA should be in equimolar amounts and a total range between 20-200 ng. For large DNA segments, increasingly proportionate amounts of DNA should be added (e.g., 250 ng of each 150 kb DNA segment). Thoroughly mix gently. Incubate the mixture at 50 °C for 15-60 min. 5 μL of the mixture is transformed into 50 μL of chemo competent *E. coli* Top 10 cells. The transformed cells are then plated on LB-agar plates that are supplemented with kanamycin.

Table 37. Master mixture for Gibson Assembly

Master mixture	
5X ISO buffer	320 μL
T5 exonuclease	0.64 μL of 10 U/ μL
Phusion polymerase	20 μL of 2 U/ μL
Taq ligase	160 μL of 40 U/ μL
dH₂O	To 2 mL
Aliquot 15μL and store at -20°C (can be stored for at least one year).	

Table 38. 5X ISO buffer for Gibson Assembly

5X ISO buffer (6 mL)	
Tris-HCl pH 7.5	3 mL of 1M stock
MgCl₂	150 μL of 2M stock
dGTP	60 μL of 100 mM stock
dATP	60 μL of 100 mM stock
dTTP	60 μL of 100 mM stock
dCTP	60 μL of 100 mM stock
DTT	300 μL of 1 M stock
PEG-8000	1.5 g
NAD	300 μL of 100 mM stock
dH₂O	To 6 mL
Aliquot 100 μL and store at -20 °C	

5.9. Preparation of chemo-competent *E. coli* cells (*E. coli* TOP10, *E. coli* BL21(DE3) and *E. coli* S-17)

For the preparation of chemo-competent *E. coli* cells, LB-medium (100 mL) was inoculated with an ONC of the desired *E. coli* strain. All media were supplemented with streptomycin (25 $\mu\text{g mL}^{-1}$). All solutions necessary were pre-cooled at 4 °C. Cultures were incubated at 37 °C and 200 rpm until an OD₆₀₀ of 0.4-0.5 was reached. Cells were harvested (4 000 x g, 15 min, 4 °C). The obtained cell pellet was resuspended in TfB1 buffer (30 mL) and MgCl₂ solution (1 M, 3.2 mL) and incubated on ice for 15 min. Cells were centrifuged (4 000 x g, 10 min, 4 °C) and the supernatant was discarded before the pellet was resuspended in TfB2 buffer (4 mL). The suspension was incubated on ice for 15 min before 50 μL aliquots were prepared in 1.5 mL tubes. Aliquots were shock-frozen in liquid nitrogen and stored at -80 °C until use.

5.10. Transformation of chemo-competent *E. coli* cells

For transformation, an aliquot of *E. coli* competent cells was thawed on ice, and typically 0.5-1 μL of plasmid DNA or 5-10 μL of a whole-plasmid-amplification reaction (e.g., QuikChange, MEGAWHOP) were added to the cell suspension. The mixture was incubated on ice for 30 min before a heat shock at 42 °C for 42 s was applied. LB-SOC medium (400 μL) was added. After the transformation procedure, cell suspensions were regenerated at 37 °C and 550 rpm for at least 1 h. Suitable aliquots were spread on antibiotic-containing LB agar plates as required and incubated at 37 °C until colony formation was observed.

5.11. Conjugation into *C. necator* H16

Once positive constructs were obtained and sequenced, transferring them from *E. coli* into *C. necator* was necessary to express the enzymes. In order to perform the conjugation, the constructs were first isolated from *E. coli* TOP10 using the GeneJET Plasmid Miniprep Kit as per manufacturer instructions. The isolated plasmids were then used to transform *E. coli* S-17 cells which are capable of acting as donors for the conjugation process.

Conjugation was performed according to the protocol of Simon et al.^[234]. Transformed *E. coli* S-17 cells were grown overnight in LB media and containing 50 $\mu\text{g mL}^{-1}$ kanamycin. In addition, the *C. necator* H16 wild-type strain was also grown overnight in TSB media containing 20 $\mu\text{g mL}^{-1}$ gentamycin, 50 $\mu\text{g mL}^{-1}$ kanamycin.

After a growth equating to an optical density (OD₆₀₀) of 0.8 -1, the overnight cultures were centrifuged in a cooled 4 °C centrifuge at a speed of 3220 x g to pellet the cells, which were resuspended in 500 μL of 0.9% saline.

200 μL of the donor (*E. coli* S17) harboring constructs plasmids was mixed with resuspended *C. necator* H16 cells. Then the mixture was dropped onto an NB media plate with no antibiotics. Let the plates dry for a few minutes. The plates were then incubated for a period of 24 hours at 28 °C.

The plate contents were suspended in 3 mL of sterile 0.9% saline solution the following next day. The cell suspensions were diluted to have a range of $10^0 - 10^{-5}$. 150 μL of the dilutions were plated on TSB plates containing 20 $\mu\text{g mL}^{-1}$ gentamycin and 200 $\mu\text{g mL}^{-1}$ kanamycin for selection. The plates were incubated at 28 °C overnight. Single, well-differentiated colonies were then picked and streaked onto fresh TSB plates with the antibiotics mentioned above and incubated at 28 °C overnight. Single colonies were then screened for the correct construct using colony PCR, and the positive hits were sent to being sequenced.

5.12. Enzyme expression in *C.necator* H16

Overnight cultures of *C.necator* H16 with the constructs were grown at 28 °C in tryptic soy broth (TSB), which was supplemented with 0.6% fructose, 20 $\mu\text{g/ml}$ gentamycin, and 200 $\mu\text{g/ml}$ kanamycin. These cultures were used to inoculate 300 mL baffled shake flasks (working volume of 50 mL) to a start OD_{600} of 0.2 for heterotrophic enzyme expression. The growth conditions in the shake flasks matched those of the overnight cultures. Cultures were allowed to grow to an OD_{600} of 0.8-1 at which they were induced with 0.1 mM of isopropyl β -D-1-thiogalactoside (IPTG) as inducer and 0.01 mM FAD as a cofactor. Cultures were grown for 18-24 hours, and then the cells were harvested and used for later use.

5.13. Enzyme expression in *E.coli* BL21(DE3)

For protein expression, a single colony of freshly transformed cells was first inoculated in 5 mL of liquid LB medium containing 50 $\mu\text{g mL}^{-1}$ kanamycin overnight. The overnight culture was used to inoculate the main culture at 37 °C until the desired cell density (OD_{600}) of approximately 0.8 was reached. Overexpression was induced by adding 0.1 mM or 0.5 mM IPTG; the cultures were grown for another 20 h at 18 °C and harvested (4000 rpm, 15 min, 4 °C).

5.14. Refolding enzyme from inclusion body

Resuspend pellets obtained after sonification in 25 ml of 0.1 M potassium phosphate buffer (pH=7) contain 6 M urea, 10 mM DTT, and 1 mM EDTA. The solution was stirred at room temperature for 2h. Centrifuge 40 min, 4 °C, 15000 rpm, keep the supernatant. Concentrating supernatant from the last step until 3 mL. 3 mL of the stock solution add dropwise, with gentle stirring to 15 mL of 0.1 M potassium phosphate buffer (pH=7), containing 1 M urea, 10% glycerol, and 1 mM DTT.(all steps were performed on ice) and concentrating the samples to 4 mL. Washing membranes with distilled water (incubate 5 min with water and slop the water). Add samples to the membrane. Put the membranes with samples into 2 L 0.1 M potassium phosphate buffer without urea, containing 10% glycerol and 1 mM DTT. The solution was dialyzed 3 h, at 4 °C with stirring. Change new buffer 2 L, and leave for overnight. Measure concentration at NANO DROP; run SDS-PAGE.

5.15. Sodium Dodecyl Sulfate- Polyacrylamide Gel Electrophoresis (SDS-PAGE) analysis

SDS-PAGE analysis was done in order to separate proteins from the cell free extract as well from the insoluble fraction. The composition of the gels, buffers, and related material can be found in the chapter 1 session.

The cell free extract was obtained by centrifuging the cell culture sample from the shake flask expressions at 3220 x g for 15 minutes at 4 °C. The obtained pellet was suspended in 500 µL of potassium phosphate buffer, and this mixture was sonicated for 30 seconds at a duty cycle of 40 twice to lyse the cells. After sonication, centrifugation at 16,000 x g for 10 minutes was done at room temperature. The supernatant obtained was the cell free extract which contained the soluble protein, and the pellet, which contained the insoluble proteins of the sample, was washed in 500 µL of potassium phosphate buffer. The supernatant was then removed, and the pellet was suspended in 500 µL of potassium phosphate buffer. The resultant suspension contained the insoluble portion of the samples.

30 µL of the samples to be tested were mixed with 10 µL of a sample buffer and boiled at 95 °C for 5 minutes. The heated samples are allowed to cool and are then pipetted into the gel. The gel was allowed to run at a current of 60 mA per gel and 200 V for 30-45 minutes. The gel was then stained overnight with a solution containing either Coomassie Brilliant Blue R-250 or G-250. It was then destained with a destaining solution if staining was done with Coomassie R-250 or with water if staining with Coomassie G-250.

5.16. Enzyme purification

The cells harboring recombinant protein were resuspended in lysis buffer, as shown in Chapter 1 Table 22. The cells were disrupted by sonication. The sonicated solution was then centrifuged (12000 rpm, 4 °C, 60 min) to remove any insoluble parts. The soluble fraction was mixed with Ni-NTA (nitrilotriacetic acid) and was incubated at 4 °C for 60 min with low-speed shaking. The column was then washed by gravity flow with the wash buffer (Chapter 1 Table 22). The bound protein was eluted with the elution buffer (Chapter 1 Table 22). The fractions were collected and analyzed by SDS-PAGE to select the ones containing the target enzyme. The enzyme solution was desalted with the desalting buffer (Chapter 1 Table 22) to remove imidazole. Finally, the enzyme samples were concentrated using Amicon® Ultra-15 centrifugal filter device (50 kDa cutoff, Millipore). The final enzyme solution was aliquoted, frozen in liquid nitrogen, and stored at -20 °C.

5.17. Activity assays with the whole-cell as the catalyst

All reactions were carried out in a 2 mL glass bottle in a total volume of 1 mL of 50 mM potassium phosphate buffer, pH 7.0. The reaction mixture contained 10mM substrates (Table 30) and 10 mg/mL of the whole cells as the catalyst. The reactions were incubated at 25 °C at 400 rpm. After 48 h, the reaction mixture was extracted with an equal volume

of dichloromethane containing acetorphan or 1-decanol as an internal standard. The organic phase was subsequently dried over MgSO₄ for gas chromatography (GC) analysis.

5.18. Activity assays with purified enzyme as a catalyst

All reactions were carried out in a 2 ml glass bottle in a total volume of 1 mL of 50 mM potassium phosphate buffer, pH 7.0. The reaction mixture contained 10mM substrates (Table 30), 10 mM NADPH, and 10 μM of the purified enzyme as the catalyst. The reactions were incubated at 25 °C at 400 rpm. After 1 h, the reaction mixture was extracted with an equal volume of dichloromethane containing acetorphan or 1-decanol as an internal standard. The organic phase was subsequently dried over MgSO₄ for gas chromatography (GC) analysis.

5.19. Organic phase extraction

400 μL of the appropriate organic solvent is mixed with the sample and vortexed vigorously to maximize mass transfer from the aqueous phase (sample) into the organic phase. The tube is then centrifuged at full speed for 5 minutes in a tabletop microcentrifuge (Eppendorf, centrifuge 5430). Phase separation can be clearly seen after this step, and the organic phase is carefully removed and mixed with some powdered magnesium sulfate in a new vial. The vial is vortexed vigorously and centrifuged for 5 minutes at full speed- this step serves to remove any traces of water from the sample. The sample is then carefully removed and transferred into a clean GC vial which can be put into the GC-FID for analysis.

5.20. Gas Chromatography- Flame Ionization Detector (GC-FID) Analysis

The GC-FID (GC-2010 Plus, Shimadzu, Japan) equipped with a ZB-5 column was used to analyze samples taken from biotransformations. Samples need to be prepared by organic phase extraction before analysis, and the appropriate methods need to be developed. Serendipitously, one method worked well enough to separate all tested substrates and their products. The method parameters used are listed in Table 39. Additionally, retention times of identified substrates, products, and side products are listed in Table 40.

Table 39. Method parameters used for GC-FID analysis.

Item	Value	Units	Control
SPL1 Temperature	230	°C	
SPL1 Pressure	35.8	kPa	
Total Flow	24	mL min ⁻¹	
Purge Flow	3	mL min ⁻¹	On
Primary Pressure	714	kPa	
Column Initial Temperature^[a]	60	°C	
FID1 Temperature	319	°C	
FID1 Makeup^[b]	30	mL min ⁻¹	On
FID1 Hydrogen Flow	40	mL min ⁻¹	On
FID1 Air Flow	399	mL min ⁻¹	On

^[a] Column temperature program: 60 °C for 5 min; 10 °C min⁻¹ to 200 °C; 30 °C min⁻¹ to 300 °C, hold 2 min.

^[b] Carrier gas: N₂.

Table 40. Retention times (given in minutes) of identified compounds relevant to this work

Compound	Retention times (min)
cyclopentanone	5.40
5-valerolactone	11.4
Thioanisole	12.0
(Methylsulfinyl)cyclohexane	16.2
Bicyclo[3.2.0]hept-2-en-6-one	9.0
Hexahydro-3H-benzofuran-2-one	14.0

References

- [1] I. Parry, "Summary for policymakers," *Implement. a US Carbon Tax Challenges Debates*, pp. xxiii–xxxiii, 2015, doi: 10.4324/9781315071961-11.
- [2] J. Murray and D. King, "Peak Oil has passed," *Nature*, vol. 481, pp. 433–435, 2012.
- [3] T. Wiesenthal *et al.*, "Biofuel support policies in Europe: Lessons learnt for the long way ahead," *Renew. Sustain. Energy Rev.*, vol. 13, no. 4, pp. 789–800, 2009, doi: 10.1016/j.rser.2008.01.011.
- [4] U. K. Royal and S. Working, "Sustainable Biofuels: Prospects and Challenges."
- [5] A. Demirbas, "Comparison of transesterification methods for production of biodiesel from vegetable oils and fats," *Energy Convers. Manag.*, vol. 49, no. 1, pp. 125–130, 2008, doi: 10.1016/j.enconman.2007.05.002.
- [6] D. Murray and U. S. A. Washington D.C., "Ethanol's potential: Looking beyond corn. ," vol. 19, no. 2, pp. 1–4, 2005, [Online]. Available: <http://www.earth-policy.org/Updates/2005/Update49.htm>.
- [7] G. Sorda, M. Banse, and C. Kemfert, "An overview of biofuel policies across the world," *Energy Policy*, vol. 38, no. 11, pp. 6977–6988, 2010, doi: 10.1016/j.enpol.2010.06.066.
- [8] G. Love *et al.*, "Continuous ethanol fermentation at 45°C using *Kluyveromyces marxianus* IMB3 immobilized in calcium alginate and kissiris," *Bioprocess Eng.*, vol. 18, no. 3, pp. 187–189, 1998, doi: 10.1007/s004490050429.
- [9] P. Nigam, I. M. Banat, D. Singh, A. P. McHale, and R. Marchant, "Continuous ethanol production by thermotolerant *Kluyveromyces marxianus* IMB3 immobilized on mineral Kissiris at 45°C," *World J. Microbiol. Biotechnol.*, vol. 13, no. 3, pp. 283–288, 1997, doi: 10.1023/A:1018578806605.
- [10] D. Brady, P. Nigam, R. Marchant, and A. P. McHale, "Ethanol production at 45°C by alginate-immobilized *Kluyveromyces marxianus* IMB3 during growth on lactose-containing media," *Bioprocess Eng.*, vol. 16, no. 2, p. 101, 1997, doi: 10.1007/s004490050295.
- [11] K. Suresh, N. Kiran sree, and L. Venkateswer Rao, "Utilization of

-
- damaged sorghum and rice grains for ethanol production by simultaneous saccharification and fermentation,” *Bioresour. Technol.*, vol. 68, no. 3, pp. 301–304, 1999, doi: 10.1016/S0960-8524(98)00135-7.
- [12] A. F. Turhollow and E. O. Heady, “Large-scale ethanol production from corn and grain sorghum and improving conversion technology,” *Energy Agric.*, vol. 5, no. 4, pp. 309–316, 1986, doi: 10.1016/0167-5826(86)90029-X.
- [13] J. C. Escobar, E. S. Lora, O. J. Venturini, E. E. Yáñez, E. F. Castillo, and O. Almazan, “Biofuels: Environment, technology and food security,” *Renew. Sustain. Energy Rev.*, vol. 13, no. 6–7, pp. 1275–1287, 2009, doi: 10.1016/j.rser.2008.08.014.
- [14] V. Patil, K. Q. Tran, and H. R. Giselrød, “Towards sustainable production of biofuels from microalgae,” *Int. J. Mol. Sci.*, vol. 9, no. 7, pp. 1188–1195, 2008, doi: 10.3390/ijms9071188.
- [15] T. M. Mata, A. A. Martins, and N. S. Caetano, “Microalgae for biodiesel production and other applications: A review,” *Renew. Sustain. Energy Rev.*, vol. 14, no. 1, pp. 217–232, 2010, doi: 10.1016/j.rser.2009.07.020.
- [16] L. Brennan and P. Owende, “Biofuels from microalgae-A review of technologies for production, processing, and extractions of biofuels and co-products,” *Renew. Sustain. Energy Rev.*, vol. 14, no. 2, pp. 557–577, 2010, doi: 10.1016/j.rser.2009.10.009.
- [17] R. Slade and A. Bauen, “Micro-algae cultivation for biofuels: Cost, energy balance, environmental impacts and future prospects,” *Biomass and Bioenergy*, vol. 53, no. 0, pp. 29–38, 2013, doi: 10.1016/j.biombioe.2012.12.019.
- [18] S. Atsumi, W. Higashide, and J. C. Liao, “Direct photosynthetic recycling of carbon dioxide to isobutyraldehyde,” *Nat. Biotechnol.*, vol. 27, no. 12, pp. 1177–1180, 2009, doi: 10.1038/nbt.1586.
- [19] N. H. Tran, J. R. Bartlett, G. S. K. Kannangara, A. S. Milev, H. Volk, and M. A. Wilson, “Catalytic upgrading of biorefinery oil from micro-algae,” *Fuel*, vol. 89, no. 2, pp. 265–274, 2010, doi: 10.1016/j.fuel.2009.08.015.
- [20] A. Ishida, “Special Section,” *Sociol. Theory Methods*, vol. 29, no. 1, pp. 17–18, 2014, doi: 10.17730/praa.3.3.1245ulmr62264266.
- [21] J. Nielsen and J. D. Keasling, “Synergies between synthetic biology and
-

-
- metabolic engineering,” *Nat. Biotechnol.*, vol. 29, no. 8, pp. 693–695, 2011, doi: 10.1038/nbt.1937.
- [22] P. P. Peralta-Yahya, F. Zhang, S. B. Del Cardayre, and J. D. Keasling, “Microbial engineering for the production of advanced biofuels,” *Nature*, vol. 488, no. 7411, pp. 320–328, 2012, doi: 10.1038/nature11478.
- [23] T. Liu and C. Khosla, “Genetic engineering of escherichia coli for biofuel production,” *Annu. Rev. Genet.*, vol. 44, pp. 53–69, 2010, doi: 10.1146/annurev-genet-102209-163440.
- [24] B. N. M. Van Leeuwen, A. M. Van Der Wulp, I. Duijnste, A. J. A. Van Maris, and A. J. J. Straathof, “Fermentative production of isobutene,” *Appl. Microbiol. Biotechnol.*, vol. 93, no. 4, pp. 1377–1387, 2012, doi: 10.1007/s00253-011-3853-7.
- [25] P. Metzger and C. Largeau, “Botryococcus braunii: A rich source for hydrocarbons and related ether lipids,” *Appl. Microbiol. Biotechnol.*, vol. 66, no. 5, pp. 486–496, 2005, doi: 10.1007/s00253-004-1779-z.
- [26] N. Ladygina, E. G. Dedyukhina, and M. B. Vainshtein, “A review on microbial synthesis of hydrocarbons,” *Process Biochem.*, vol. 41, no. 5, pp. 1001–1014, 2006, doi: 10.1016/j.procbio.2005.12.007.
- [27] R. W. Howard and G. J. Blomquist, “Ecological, behavioral, and biochemical aspects of insect hydrocarbons,” *Annu. Rev. Entomol.*, vol. 50, pp. 371–393, 2005, doi: 10.1146/annurev.ento.50.071803.130359.
- [28] A. Bernard and J. Joubès, “Arabidopsis cuticular waxes: Advances in synthesis, export and regulation,” *Prog. Lipid Res.*, vol. 52, no. 1, pp. 110–129, 2013, doi: 10.1016/j.plipres.2012.10.002.
- [29] M. G. Aarts, C. J. Keijzer, W. J. Stiekema, and A. Pereira, “Molecular characterization of the CER1 gene of Arabidopsis involved in epicuticular wax biosynthesis and pollen fertility,” *Plant Cell*, vol. 7, no. 12, pp. 2115–2127, 1995, doi: 10.1105/tpc.7.12.2115.
- [30] J. D. Weete, “Aliphatic hydrocarbons of the fungi,” *Phytochemistry*, vol. 11, no. 4, pp. 1201–1205, 1972, doi: 10.1016/S0031-9422(00)90066-9.
- [31] J. Wang and K. Zhu, “Microbial production of alka(e)ne biofuels,” *Curr. Opin. Biotechnol.*, vol. 50, pp. 11–18, 2018, doi: 10.1016/j.copbio.2017.08.009.
- [32] D. J. Sukovich, J. L. Seffernick, J. E. Richman, J. A. Gralnick, and L. P. Wackett, “Widespread head-to-head hydrocarbon biosynthesis in
-

-
- bacteria and role of OleA,” *Appl. Environ. Microbiol.*, vol. 76, no. 12, pp. 3850–3862, 2010, doi: 10.1128/AEM.00436-10.
- [33] J. A. Frias, J. E. Richman, J. S. Erickson, and L. P. Wackett, “Purification and characterization of OleA from *Xanthomonas campestris* and demonstration of a non-decarboxylative claisen condensation reaction,” *J. Biol. Chem.*, vol. 286, no. 13, pp. 10930–10938, 2011, doi: 10.1074/jbc.M110.216127.
- [34] J. A. Frias, B. R. Goblirsch, L. P. Wackett, and C. M. Wilmot, “Cloning, purification, crystallization and preliminary X-ray diffraction of the OleC protein from *Stenotrophomonas maltophilia* involved in head-to-head hydrocarbon biosynthesis,” *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.*, vol. 66, no. 9, pp. 1108–1110, 2010, doi: 10.1107/S1744309110031751.
- [35] D. Mendez-Perez, M. B. Begemann, and B. F. Pfeleger, “Modular synthase-encoding gene involved in α -olefin biosynthesis in *Synechococcus* sp. strain pcC 7002,” *Appl. Environ. Microbiol.*, vol. 77, no. 12, pp. 4264–4267, 2011, doi: 10.1128/AEM.00467-11.
- [36] S. Smith and S. C. Tsai, “The type I fatty acid and polyketide synthases: A tale of two megasynthases,” *Nat. Prod. Rep.*, vol. 24, no. 5, pp. 1041–1072, 2007, doi: 10.1039/b603600g.
- [37] C. J. Knoot and H. B. Pakrasi, “Diverse hydrocarbon biosynthetic enzymes can substitute for olefin synthase in the cyanobacterium *Synechococcus* sp. PCC 7002,” *Sci. Rep.*, vol. 9, no. 1, pp. 1–12, 2019, doi: 10.1038/s41598-018-38124-y.
- [38] M. A. Rude, T. S. Baron, S. Brubaker, M. Alibhai, S. B. Del Cardayre, and A. Schirmer, “Terminal olefin (1-alkene) biosynthesis by a novel P450 fatty acid decarboxylase from *Jeotgalicoccus* species,” *Appl. Environ. Microbiol.*, vol. 77, no. 5, pp. 1718–1727, 2011, doi: 10.1128/AEM.02580-10.
- [39] Y. Liu *et al.*, “Hydrogen peroxide-independent production of α -alkenes by OleT JE P450 fatty acid decarboxylase,” *Biotechnol. Biofuels*, vol. 7, no. 1, pp. 1–12, 2014, doi: 10.1186/1754-6834-7-28.
- [40] Z. Rui *et al.*, “Microbial biosynthesis of medium-chain 1-alkenes by a nonheme iron oxidase,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 111, no. 51, pp. 18237–18242, 2014, doi: 10.1073/pnas.1419701112.
- [41] Z. Rui, N. C. Harris, X. Zhu, W. Huang, and W. Zhang, “Discovery of a

-
- Family of Desaturase-Like Enzymes for 1-Alkene Biosynthesis,” *ACS Catal.*, vol. 5, no. 12, pp. 7091–7094, 2015, doi: 10.1021/acscatal.5b01842.
- [42] A. Bernard *et al.*, “Reconstitution of plant alkane biosynthesis in yeast demonstrates that *Arabidopsis* ECERIFERUM1 and ECERIFERUM3 are core components of a very-long-chain alkane synthesis complex,” *Plant Cell*, vol. 24, no. 7, pp. 3106–3118, 2012, doi: 10.1105/tpc.112.099796.
- [43] Y. Qiu *et al.*, “An insect-specific P450 oxidative decarboxylase for cuticular hydrocarbon biosynthesis,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 109, no. 37, pp. 14858–14863, 2012, doi: 10.1073/pnas.1208650109.
- [44] A. M. Kunjapur and K. L. J. Prather, “Microbial engineering for aldehyde synthesis,” *Appl. Environ. Microbiol.*, vol. 81, no. 6, pp. 1892–1901, 2015, doi: 10.1128/AEM.03319-14.
- [45] L. J. Rajakovich *et al.*, “Rapid Reduction of the Diferric-Peroxyhemiacetal Intermediate in Aldehyde-Deformylating Oxygenase by a Cyanobacterial Ferredoxin: Evidence for a Free-Radical Mechanism,” *J. Am. Chem. Soc.*, vol. 137, no. 36, pp. 11695–11709, 2015, doi: 10.1021/jacs.5b06345.
- [46] D. Sorigué *et al.*, “An algal photoenzyme converts fatty acids to hydrocarbons,” *Science (80-.)*, vol. 357, no. 6354, pp. 903–907, 2017, doi: 10.1126/science.aan6349.
- [47] D. Sorigué *et al.*, “Microalgae synthesize hydrocarbons from long-chain fatty acids via a light-dependent pathway,” *Plant Physiol.*, vol. 171, no. 4, pp. 2393–2405, 2016, doi: 10.1104/pp.16.00462.
- [48] A. Benjdia, “DNA photolyases and SP lyase: Structure and mechanism of light-dependent and independent DNA lyases,” *Curr. Opin. Struct. Biol.*, vol. 22, no. 6, pp. 711–720, 2012, doi: 10.1016/j.sbi.2012.10.002.
- [49] K. Brettel and M. Byrdin, “Reaction mechanisms of DNA photolyase,” *Curr. Opin. Struct. Biol.*, vol. 20, no. 6, pp. 693–701, 2010, doi: 10.1016/j.sbi.2010.07.003.
- [50] D. Zhong, “Electron transfer mechanisms of DNA repair by photolyase,” *Annu. Rev. Phys. Chem.*, vol. 66, pp. 691–715, 2015, doi: 10.1146/annurev-physchem-040513-103631.
- [51] A. Sancar, “Mechanisms of DNA Repair by Photolyase and Excision
-

-
- Nuclease (Nobel Lecture),” *Angew. Chemie - Int. Ed.*, vol. 55, no. 30, pp. 8502–8527, 2016, doi: 10.1002/anie.201601524.
- [52] H. M. Wilks and M. P. Timko, “A light-dependent complementation system for analysis of NADPH:protochlorophyllide oxidoreductase: Identification and mutagenesis of two conserved residues that are essential for enzyme activity,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 92, no. 3, pp. 724–728, 1995, doi: 10.1073/pnas.92.3.724.
- [53] M. Y. Ho, G. Shen, D. P. Canniffe, C. Zhao, and D. A. Bryant, “Light-dependent chlorophyll f synthase is a highly divergent paralog of PsbA of photosystem II,” *Science (80-.)*, vol. 353, no. 6302, 2016, doi: 10.1126/science.aaf9178.
- [54] “Mechanism and dynamics of fatty acid photodecarboxylase,” vol. 148, no. April, 2021, doi: 10.1126/science.abd5687.
- [55] D. J. Heyes, B. Lakavath, S. J. O. Hardman, M. Sakuma, T. M. Hedison, and N. S. Scrutton, “Photochemical Mechanism of Light-Driven Fatty Acid Photodecarboxylase,” *ACS Catal.*, vol. 10, no. 12, pp. 6691–6696, 2020, doi: 10.1021/acscatal.0c01684.
- [56] B. Lakavath *et al.*, “Radical-based photoinactivation of fatty acid photodecarboxylases,” *Anal. Biochem.*, vol. 600, no. March, p. 113749, 2020, doi: 10.1016/j.ab.2020.113749.
- [57] M. M. E. Huijbers, W. Zhang, F. Tonin, and F. Hollmann, “Light-Driven Enzymatic Decarboxylation of Fatty Acids,” *Angew. Chemie - Int. Ed.*, vol. 57, no. 41, pp. 13648–13651, 2018, doi: 10.1002/anie.201807119.
- [58] R. Hoeven *et al.*, “Distributed Biomanufacturing of Liquefied Petroleum Gas,” *bioRxiv*, p. 640474, 2019, doi: 10.1101/640474.
- [59] M. Amer *et al.*, “Low carbon strategies for sustainable bio-alkane gas production and renewable energy,” *Energy Environ. Sci.*, vol. 13, no. 6, pp. 1818–1831, 2020, doi: 10.1039/d0ee00095g.
- [60] M. Amer *et al.*, “Renewable and tuneable bio-LPG blends derived from amino acids,” *Biotechnol. Biofuels*, vol. 13, no. 1, pp. 1–15, 2020, doi: 10.1186/s13068-020-01766-0.
- [61] N. Menon *et al.*, “A microbial platform for renewable propane synthesis based on a fermentative butanol pathway,” *Biotechnol. Biofuels*, vol. 8, no. 1, pp. 1–12, 2015, doi: 10.1186/s13068-015-0231-1.
- [62] D. Gocke *et al.*, “Comparative characterisation of thiamin diphosphate-
-

-
- dependent decarboxylases,” *J. Mol. Catal. B Enzym.*, vol. 61, no. 1–2, pp. 30–35, 2009, doi: 10.1016/j.molcatb.2009.03.019.
- [63] W. Zhang *et al.*, “Photobiocatalytic synthesis of chiral secondary fatty alcohols from renewable unsaturated fatty acids,” *Nat. Commun.*, vol. 11, no. 1, pp. 1–8, 2020, doi: 10.1038/s41467-020-16099-7.
- [64] J. Xu *et al.*, “Light-Driven Kinetic Resolution of α -Functionalized Carboxylic Acids Enabled by an Engineered Fatty Acid Photodecarboxylase,” *Angew. Chemie - Int. Ed.*, vol. 58, no. 25, pp. 8474–8478, 2019, doi: 10.1002/anie.201903165.
- [65] F. Cheng *et al.*, “Light-driven deracemization of phosphinothricin by engineered fatty acid photodecarboxylase on a gram scale,” *Green Chem.*, vol. 22, no. 20, pp. 6815–6818, 2020, doi: 10.1039/d0gc02696d.
- [66] P. Xua, K. Qiao, W. S. Ahn, and G. Stephanopoulos, “Engineering *Yarrowia lipolytica* as a platform for synthesis of drop-in transportation fuels and oleochemicals,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 113, no. 39, pp. 10848–10853, 2016, doi: 10.1073/pnas.1607295113.
- [67] J. Li *et al.*, “Synthesis of high-titer alka(e)nes in *Yarrowia lipolytica* is enabled by a discovered mechanism,” *Nat. Commun.*, vol. 11, no. 1, 2020, doi: 10.1038/s41467-020-19995-0.
- [68] M. Karava, P. Gockel, and J. Kabisch, “*Bacillus subtilis* spore surface display of photodecarboxylase for the transformation of lipids to hydrocarbons,” *bioRxiv*, p. 2020.08.30.273821, 2020, [Online]. Available: <https://doi.org/10.1101/2020.08.30.273821>.
- [69] W. Zhang *et al.*, “Hydrocarbon Synthesis via Photoenzymatic Decarboxylation of Carboxylic Acids,” *J. Am. Chem. Soc.*, vol. 141, no. 7, pp. 3116–3120, 2019, doi: 10.1021/jacs.8b12282.
- [70] M. J. Harms and J. W. Thornton, “Analyzing protein structure and function using ancestral gene reconstruction,” *Curr. Opin. Struct. Biol.*, vol. 20, no. 3, pp. 360–366, 2010, doi: 10.1016/j.sbi.2010.03.005.
- [71] R. Merkl and R. Sterner, “Ancestral protein reconstruction: Techniques and applications,” *Biol. Chem.*, vol. 397, no. 1, pp. 1–21, 2016, doi: 10.1515/hsz-2015-0158.
- [72] M. Wilding, T. S. Peat, S. Kalyaanamoorthy, J. Newman, C. Scott, and L. S. Jermini, “Reverse engineering: Transaminase biocatalyst development using ancestral sequence reconstruction,” *Green Chem.*,
-

-
- vol. 19, no. 22, pp. 5375–5380, 2017, doi: 10.1039/c7gc02343j.
- [73] D. L. Trudeau, M. Kaltenbach, and D. S. Tawfik, “On the Potential Origins of the High Stability of Reconstructed Ancestral Proteins,” *Mol. Biol. Evol.*, vol. 33, no. 10, pp. 2633–2641, 2016, doi: 10.1093/molbev/msw138.
- [74] A. Thomas, R. Cutlan, W. Finnigan, M. van der Giezen, and N. Harmer, “Highly thermostable carboxylic acid reductases generated by ancestral sequence reconstruction,” *Commun. Biol.*, vol. 2, no. 1, pp. 1–12, 2019, doi: 10.1038/s42003-019-0677-y.
- [75] J. C. Francis and P. E. Hansche, “Directed evolution of metabolic pathways in microbial populations. I. Modification of the acid phosphatase pH optimum in *S. cerevisiae*.” *Genetics*, vol. 70, no. 1, pp. 59–73, 1972, doi: 10.1093/genetics/70.1.59.
- [76] H. W. Hellinga, “Rational protein design: Combining theory and experiment,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 94, no. 19, pp. 10015–10017, 1997, doi: 10.1073/pnas.94.19.10015.
- [77] S. Lutz, “Beyond directed evolution-semi-rational protein engineering and design,” *Curr. Opin. Biotechnol.*, vol. 21, no. 6, pp. 734–743, 2010, doi: 10.1016/j.copbio.2010.08.011.
- [78] C. K. Savile *et al.*, “Biocatalytic Asymmetric Synthesis of Sitagliptin Manufacture,” *Science (80-)*, vol. 329, no. July, pp. 305–310, 2010, doi: 10.1126/science.1188934.
- [79] P. Heinzelman *et al.*, “A family of thermostable fungal cellulases created by structure-guided recombination,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 106, no. 14, pp. 5610–5615, 2009, doi: 10.1073/pnas.0901417106.
- [80] J. Damborsky and J. Brezovsky, “Computational tools for designing and engineering enzymes,” *Curr. Opin. Chem. Biol.*, vol. 19, no. 1, pp. 8–16, 2014, doi: 10.1016/j.cbpa.2013.12.003.
- [81] P. Larrañaga *et al.*, “Machine learning in bioinformatics,” *Brief. Bioinform.*, vol. 7, no. 1, pp. 86–112, 2006, doi: 10.1093/bib/bbk007.
- [82] P. Jones *et al.*, “InterProScan 5: Genome-scale protein function classification,” *Bioinformatics*, vol. 30, no. 9, pp. 1236–1240, 2014, doi: 10.1093/bioinformatics/btu031.
- [83] M. Musil *et al.*, “FireProt: web server for automated design of thermostable proteins,” *Nucleic Acids Res.*, vol. 45, no. W1, pp. W393–
-

-
- W399, 2017, doi: 10.1093/nar/gkx285.
- [84] L. Sumbalova, J. Stourac, T. Martinek, D. Bednar, and J. Damborsky, "HotSpot Wizard 3.0: Web server for automated design of mutations and smart libraries based on sequence input information," *Nucleic Acids Res.*, vol. 46, no. W1, pp. W356–W362, 2018, doi: 10.1093/nar/gky417.
- [85] D. R. Canchi and A. E. García, "Cosolvent effects on protein stability," *Annu. Rev. Phys. Chem.*, vol. 64, pp. 273–293, 2013, doi: 10.1146/annurev-physchem-040412-110156.
- [86] R. Verma and T. Banerjee, "Liquid-Liquid Extraction of Lower Alcohols Using Menthol-Based Hydrophobic Deep Eutectic Solvent: Experiments and COSMO-SAC Predictions," *Ind. Eng. Chem. Res.*, vol. 57, no. 9, pp. 3371–3381, 2018, doi: 10.1021/acs.iecr.7b05270.
- [87] A. Paiva, R. Craveiro, I. Aroso, M. Martins, R. L. Reis, and A. R. C. Duarte, "Natural deep eutectic solvents - Solvents for the 21st century," *ACS Sustain. Chem. Eng.*, vol. 2, no. 5, pp. 1063–1071, 2014, doi: 10.1021/sc500096j.
- [88] E. R. Parnham, E. A. Drylie, P. S. Wheatley, A. M. Z. Slawin, and R. E. Morris, "Ionothermal Materials Synthesis Using Unstable Deep-Eutectic Solvents as Template-Delivery Agents," *Angew. Chemie*, vol. 118, no. 30, pp. 5084–5088, 2006, doi: 10.1002/ange.200600290.
- [89] A. P. Abbott, G. Capper, D. L. Davies, R. K. Rasheed, and V. Tambyrajah, "Novel solvent properties of choline chloride/urea mixtures," *Chem. Commun.*, no. 1, pp. 70–71, 2003, doi: 10.1039/b210714g.
- [90] Q. Zhang, K. De Oliveira Vigier, S. Royer, and F. Jérôme, "Deep eutectic solvents: Syntheses, properties and applications," *Chem. Soc. Rev.*, vol. 41, no. 21, pp. 7108–7146, 2012, doi: 10.1039/c2cs35178a.
- [91] J. T. Gorke, F. Sreenc, and R. J. Kazlauskas, "Hydrolase-catalyzed biotransformations in deep eutectic solvents," *Chem. Commun.*, no. 10, pp. 1235–1237, 2008, doi: 10.1039/b716317g.
- [92] A. K. Schweiger *et al.*, "Using Deep Eutectic Solvents to Overcome Limited Substrate Solubility in the Enzymatic Decarboxylation of Bio-Based Phenolic Acids," *ACS Sustain. Chem. Eng.*, 2019, doi: 10.1021/acssuschemeng.9b03455.
- [93] M. C. Gutiérrez, M. L. Ferrer, L. Yuste, F. Rojo, and F. Del Monte,
-

-
- “Bacteria incorporation in deep-eutectic solvents through freeze-drying,” *Angew. Chemie - Int. Ed.*, vol. 49, no. 12, pp. 2158–2162, 2010, doi: 10.1002/anie.200905212.
- [94] K. Kimura, M. Yamaoka, and Y. Kamisaka, “Rapid estimation of lipids in oleaginous fungi and yeasts using Nile red fluorescence,” *J. Microbiol. Methods*, vol. 56, no. 3, pp. 331–338, 2004, doi: 10.1016/j.mimet.2003.10.018.
- [95] P. Greenspan and S. D. Fowler, “Spectrofluorometric studies of the lipid probe, Nile red,” *J. Lipid Res.*, vol. 26, no. 7, pp. 781–789, 1985, doi: 10.1016/s0022-2275(20)34307-8.
- [96] D. L. Sackett and J. Wolff, “Nile red as a polarity-sensitive fluorescent probe of hydrophobic protein surfaces,” *Anal. Biochem.*, vol. 167, no. 2, pp. 228–234, 1987, doi: 10.1016/0003-2697(87)90157-6.
- [97] P. Parthasarathy and S. K. Narayanan, “Effect of Hydrothermal Carbonization Reaction Parameters on,” *Environ. Prog. Sustain. Energy*, vol. 33, no. 3, pp. 676–680, 2014, doi: 10.1002/ep.
- [98] P. Greenspan, E. P. Mayer, and S. D. Fowler, “Nile red: A selective fluorescent stain for intracellular lipid droplets,” *J. Cell Biol.*, vol. 100, no. 3, pp. 965–973, 1985, doi: 10.1083/jcb.100.3.965.
- [99] V. Gorenflo, A. Steinbüchel, S. Marose, M. Rieseberg, and T. Scheper, “Quantification of bacterial polyhydroxyalkanoic acids by Nile red staining,” *Appl. Microbiol. Biotechnol.*, vol. 51, no. 6, pp. 765–772, 1999, doi: 10.1007/s002530051460.
- [100] W. Chen, C. Zhang, L. Song, M. Sommerfeld, and Q. Hu, “A high throughput Nile red method for quantitative measurement of neutral lipids in microalgae,” *J. Microbiol. Methods*, vol. 77, no. 1, pp. 41–47, 2009, doi: 10.1016/j.mimet.2009.01.001.
- [101] A. Leaver-Fay *et al.*, “Rosetta3: An object-oriented software suite for the simulation and design of macromolecules,” *Methods Enzymol.*, vol. 487, no. C, pp. 545–574, 2011, doi: 10.1016/B978-0-12-381270-4.00019-6.
- [102] G. A. Broadwell, “Valency, transitivity, and passive constructions in Kaqchikel,” *Biotechnol. Lett.*, pp. 1385–1395, 2003, doi: 10.1023/A:1025024104862.
- [103] D. M. Francis and R. Page, “Strategies to optimize protein expression in *E. coli*,” *Curr. Protoc. Protein Sci.*, no. SUPPL. 61, 2010, doi:
-

10.1002/0471140864.ps0524s61.

- [104]G. Despotou, I. Korkontzelos, N. Matragkas, E. Bilici, and T. N. Arvanitis, “Structuring Clinical Decision Support Rules for Drug Safety Using Natural Language Processing,” *Stud. Health Technol. Inform.*, vol. 251, no. February, pp. 89–92, 2018, doi: 10.3233/978-1-61499-880-8-89.
- [105]P. Malakar and K. V. Venkatesh, “Effect of substrate and IPTG concentrations on the burden to growth of *Escherichia coli* on glycerol due to the expression of Lac proteins,” *Appl. Microbiol. Biotechnol.*, vol. 93, no. 6, pp. 2543–2549, 2012, doi: 10.1007/s00253-011-3642-3.
- [106]R. S. Donovan, C. W. Robinson, and B. R. Click, “Review: Optimizing inducer and culture conditions for expression of foreign proteins under the control of the lac promoter,” *J. Ind. Microbiol.*, vol. 16, no. 3, pp. 145–154, 1996, doi: 10.1007/BF01569997.
- [107]D. J. Wurm *et al.*, “Teaching an old pET new tricks: tuning of inclusion body formation and properties by a mixed feed system in *E. coli*,” *Appl. Microbiol. Biotechnol.*, vol. 102, no. 2, pp. 667–676, 2018, doi: 10.1007/s00253-017-8641-6.
- [108]F. T. Haddadin and S. W. Harcum, “Transcriptome profiles for high-cell-density recombinant and wild-type *Escherichia coli*,” *Biotechnol. Bioeng.*, vol. 90, no. 2, pp. 127–153, 2005, doi: 10.1002/bit.20340.
- [109]R. Kazlauskas, “Engineering more stable proteins,” *Chem. Soc. Rev.*, vol. 47, no. 24, pp. 9026–9045, 2018, doi: 10.1039/c8cs00014j.
- [110]F. Forneris, R. Orru, D. Bonivento, L. R. Chiarelli, and A. Mattevi, “ThermoFAD, a Thermofluor®-adapted flavin ad hoc detection system for protein folding and ligand binding,” *FEBS J.*, vol. 276, no. 10, pp. 2833–2840, 2009, doi: 10.1111/j.1742-4658.2009.07006.x.
- [111]C. Yin *et al.*, “An algal photoenzyme converts fatty acids to hydrocarbons,” *Science (80-.)*, vol. 100190, no. 7, pp. 965–973, 2018, doi: 10.1126/science.aan6349.
- [112]J. Yang, I. Anishchenko, H. Park, Z. Peng, S. Ovchinnikov, and D. Baker, “Improved protein structure prediction using predicted interresidue orientations,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 117, no. 3, pp. 1496–1503, 2020, doi: 10.1073/pnas.1914677117.
- [113]Y. Vander Meersche, G. Cretin, A. G. de Brevern, J.-C. Gelly, and T. Galochkina, “MEDUSA: Prediction of Protein Flexibility from Sequence,”

-
- J. Mol. Biol.*, no. December 2020, p. 166882, 2021, doi: 10.1016/j.jmb.2021.166882.
- [114]Y. Ma *et al.*, “Photoenzymatic Production of Next Generation Biofuels from Natural Triglycerides Combining a Hydrolase and a Photodecarboxylase,” *ChemPhotoChem*, vol. 4, no. 1, pp. 39–44, 2020, doi: 10.1002/cptc.201900205.
- [115]C. C. C. R. De Carvalho, “Enzymatic and whole cell catalysis: Finding new strategies for old processes,” *Biotechnol. Adv.*, vol. 29, no. 1, pp. 75–83, 2011, doi: 10.1016/j.biotechadv.2010.09.001.
- [116]R. Xin, S. Qi, C. Zeng, F. I. Khan, B. Yang, and Y. Wang, “A functional natural deep eutectic solvent based on trehalose: Structural and physicochemical properties,” *Food Chem.*, vol. 217, pp. 560–567, 2017, doi: 10.1016/j.foodchem.2016.09.012.
- [117]S. K. Shukla and J. P. Mikkola, “Intermolecular interactions upon carbon dioxide capture in deep-eutectic solvents,” *Phys. Chem. Chem. Phys.*, vol. 20, no. 38, pp. 24591–24601, 2018, doi: 10.1039/c8cp03724h.
- [118]D. J. G. P. Van Osch, L. F. Zubeir, A. Van Den Bruinhorst, M. A. A. Rocha, and M. C. Kroon, “Hydrophobic deep eutectic solvents as water-immiscible extractants,” *Green Chem.*, vol. 17, no. 9, pp. 4518–4521, 2015, doi: 10.1039/c5gc01451d.
- [119]C. Florindo, L. Romero, I. Rintoul, L. C. Branco, and I. M. Marrucho, “From Phase Change Materials to Green Solvents: Hydrophobic Low Viscous Fatty Acid-Based Deep Eutectic Solvents,” *ACS Sustain. Chem. Eng.*, vol. 6, no. 3, pp. 3888–3895, 2018, doi: 10.1021/acssuschemeng.7b04235.
- [120]B. D. Ribeiro, C. Florindo, L. C. Iff, M. A. Z. Coelho, and I. M. Marrucho, “Menthol-based eutectic mixtures: Hydrophobic low viscosity solvents,” *ACS Sustain. Chem. Eng.*, vol. 3, no. 10, pp. 2469–2477, 2015, doi: 10.1021/acssuschemeng.5b00532.
- [121]C. Florindo, L. C. Branco, and I. M. Marrucho, “Quest for Green-Solvent Design: From Hydrophilic to Hydrophobic (Deep) Eutectic Solvents,” *ChemSusChem*, vol. 12, no. 8, pp. 1549–1559, 2019, doi: 10.1002/cssc.201900147.
- [122]W. Zhang *et al.*, “Hydrocarbon Synthesis via Photoenzymatic Decarboxylation of Carboxylic Acids,” *J. Am. Chem. Soc.*, vol. 141, no. 7, pp. 3116–3120, 2019, doi: 10.1021/jacs.8b12282.
-

-
- [123]F. Sievers *et al.*, “Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega,” *Mol. Syst. Biol.*, vol. 7, no. 539, 2011, doi: 10.1038/msb.2011.75.
- [124]Y. Gumulya *et al.*, “Engineering highly functional thermostable proteins using ancestral sequence reconstruction,” *Nat. Catal.*, vol. 1, no. 11, pp. 878–888, 2018, doi: 10.1038/s41929-018-0159-5.
- [125]C. Notredame, D. G. Higgins, and J. Heringa, “T-coffee: A novel method for fast and accurate multiple sequence alignment,” *J. Mol. Biol.*, vol. 302, no. 1, pp. 205–217, 2000, doi: 10.1006/jmbi.2000.4042.
- [126]D. T. Jones, W. R. Taylor, and J. M. Thornton, “The rapid generation of mutation data matrices from protein sequences,” *Bioinformatics*, vol. 8, no. 3, pp. 275–282, 1992, doi: 10.1093/bioinformatics/8.3.275.
- [127]S. Q. Le and O. Gascuel, “An improved general amino acid replacement matrix,” *Mol. Biol. Evol.*, vol. 25, no. 7, pp. 1307–1320, 2008, doi: 10.1093/molbev/msn067.
- [128]K. Katoh, K. Misawa, K. I. Kuma, and T. Miyata, “MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform,” *Nucleic Acids Res.*, vol. 30, no. 14, pp. 3059–3066, 2002, doi: 10.1093/nar/gkf436.
- [129]Schrodinger LLC, “The AxPyMOL Molecular Graphics Plugin for Microsoft PowerPoint, Version 1.8.” 2015.
- [130]E. Gasteiger *et al.*, “Protein Identification and Analysis Tools on the ExPASy Server,” in *The Proteomics Protocols Handbook*, J. M. Walker, Ed. Totowa, NJ: Humana Press, 2005, pp. 571–607.
- [131]B. Kuhlman, G. Dantas, G. C. Ireton, G. Varani, B. L. Stoddard, and D. Baker, “Atomic-Level Accuracy,” vol. 302, no. November, pp. 1364–1369, 2003.
- [132]S. Harayama, M. Kok, and E. L. Neidle, “Functional and evolutionary relationships among diverse oxygenases,” *Annu. Rev. Microbiol.*, vol. 46, pp. 565–601, 1992, doi: 10.1146/annurev.mi.46.100192.003025.
- [133]J. B. Van Beilen and E. G. Funhoff, “Expanding the alkane oxygenase toolbox: New enzymes and applications,” *Curr. Opin. Biotechnol.*, vol. 16, no. 3 SPEC. ISS., pp. 308–314, 2005, doi: 10.1016/j.copbio.2005.04.005.
- [134]J. Chakraborty, D. Ghosal, A. Dutta, and T. K. Dutta, “An insight into the

-
- origin and functional evolution of bacterial aromatic ring-hydroxylating oxygenases,” *J. Biomol. Struct. Dyn.*, vol. 30, no. 4, pp. 419–436, 2012, doi: 10.1080/07391102.2012.682208.
- [135]A. J. Mitchell and J. K. Weng, “Unleashing the synthetic power of plant oxygenases: From mechanism to application,” *Plant Physiol.*, vol. 179, no. 3, pp. 813–829, 2019, doi: 10.1104/pp.18.01223.
- [136]J. B. Van Beilen, W. A. Duetz, A. Schmid, and B. Witholt, “Practical issues in the application of oxygenases,” *Trends Biotechnol.*, vol. 21, no. 4, pp. 170–177, 2003, doi: 10.1016/S0167-7799(03)00032-5.
- [137]S. G. Burton, “Oxidizing enzymes as biocatalysts,” *Trends Biotechnol.*, vol. 21, no. 12, pp. 543–549, 2003, doi: 10.1016/j.tibtech.2003.10.006.
- [138]J. B. Bonanno *et al.*, “Mechanism of action of a flavin- containing monooxygenase,” vol. 104, no. 36, 2007.
- [139]V. B. Urlacher and S. Eiben, “Cytochrome P450 monooxygenases: perspectives for synthetic application,” *Trends Biotechnol.*, vol. 24, no. 7, pp. 324–330, 2006, doi: 10.1016/j.tibtech.2006.05.002.
- [140]T. D. H. Bugg, “Dioxygenase enzymes: Catalytic mechanisms and chemical models,” *Tetrahedron*, vol. 59, no. 36, pp. 7075–7101, 2003, doi: 10.1016/S0040-4020(03)00944-X.
- [141]J. Olucha and A. L. Lamb, “Mechanistic and structural studies of the N-hydroxylating flavoprotein monooxygenases,” *Bioorg. Chem.*, vol. 39, no. 5–6, pp. 171–177, 2011, doi: 10.1016/j.bioorg.2011.07.006.
- [142]N. L. Schlaich, “Flavin-containing monooxygenases in plants: looking beyond detox,” *Trends Plant Sci.*, vol. 12, no. 9, pp. 412–418, 2007, doi: 10.1016/j.tplants.2007.08.009.
- [143]P. S. Phale, A. Basu, P. D. Majhi, J. Deveryshetty, C. Vamsee-Krishna, and R. Shrivastava, “Metabolic diversity in bacterial degradation of aromatic compounds,” *Omi. A J. Integr. Biol.*, vol. 11, no. 3, pp. 252–279, 2007, doi: 10.1089/omi.2007.0004.
- [144]M. Ahmad, J. N. Roberts, E. M. Hardiman, R. Singh, L. D. Eltis, and T. D. H. Bugg, “Identification of DypB from *rhodococcus jostii* RHA1 as a lignin peroxidase,” *Biochemistry*, vol. 50, no. 23, pp. 5096–5107, 2011, doi: 10.1021/bi101892z.
- [145]P. Kallio, Z. Liu, P. Mäntsälä, J. Niemi, and M. Metsä-Ketelä, “Sequential Action of Two Flavoenzymes, PgaE and PgaM, in
-

-
- Angucycline Biosynthesis: Chemoenzymatic Synthesis of Gaudimycin C,” *Chem. Biol.*, vol. 15, no. 2, pp. 157–166, 2008, doi: 10.1016/j.chembiol.2007.12.011.
- [146]K. S. Ryan, A. R. Howard-Jones, M. J. Hamill, S. J. Elliott, C. T. Walsh, and C. L. Drennan, “Crystallographic trapping in the rebeccamycin biosynthetic enzyme RebC,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 104, no. 39, pp. 15311–15316, 2007, doi: 10.1073/pnas.0707190104.
- [147]A. J. Fisher, I. Rayment, F. M. Raushel, and T. O. Baldwin, “Three-Dimensional Structure of Bacterial Luciferase from *Vibrio Harveyi* at 2.4 Å Resolution,” *Biochemistry*, vol. 34, no. 20, pp. 6581–6586, 1995, doi: 10.1021/bi00020a002.
- [148]D. Tischler, D. Eulberg, S. Lakner, S. R. Kaschabek, W. J. H. Van Berkel, and M. Schlömann, “Identification of a novel self-sufficient styrene monooxygenase from *Rhodococcus opacus* 1CP,” *J. Bacteriol.*, vol. 191, no. 15, pp. 4996–5009, 2009, doi: 10.1128/JB.00307-09.
- [149]R. Systems, *Handbook of Flavoproteins*. 2013.
- [150]Adolf Baeyer and Victor Villiger, “Einwirkung den Caro’sohen Reagens auf Ketone.,” *Zeitschr. f. angewandte Chemie*, 1898.
- [151]H. Leisch, K. Morley, and P. C. K. Lau, “Baeyer-villiger monooxygenases: More than just green chemistry,” *Chem. Rev.*, vol. 111, no. 7, pp. 4165–4222, 2011, doi: 10.1021/cr1003437.
- [152]R. Bruckner, M. Harmata, and P. A. Wender, “Organic mechanisms: Reactions, stereochemistry and synthesis,” *Organic Mechanisms: Reactions, Stereochemistry and Synthesis*. pp. 1–855, 2010, doi: 10.1007/978-3-642-03651-4.
- [153]D. Hingerty and O. Reilly, “Microbiological degradation of sterols,” *Nature*, vol. 152, no. 3855, p. 332, 1943.
- [154]P. W. Trudgill, “Microbial metabolism and transformation of selected monoterpenes,” *Biochem. Microb. Degrad.*, pp. 33–61, 1994, doi: 10.1007/978-94-011-1687-9_2.
- [155]F. Secundo, G. Carrea, S. Dallavalle, and G. Franzosi, “Asymmetric oxidation of sulfides by cyclohexanone monooxygenase,” *Tetrahedron: Asymmetry*, vol. 4, no. 9, pp. 1981–1982, 1993, doi: 10.1016/S0957-4166(00)82244-2.
- [156]S. Colonna, N. Gaggero, P. Pasta, and G. Ottolina, “Enantioselective

-
- oxidation of sulfides to sulfoxides catalysed by bacterial cyclohexanone monooxygenases,” *Chem. Commun.*, no. 20, pp. 2303–2307, 1996, doi: 10.1039/cc9960002303.
- [157]G. Ottolina, S. Bianchi, B. Belloni, G. Carrea, and B. Danieli, “First asymmetric oxidation of tertiary amines by cyclohexanone monooxygenase,” *Tetrahedron Lett.*, vol. 40, no. 48, pp. 8483–8486, 1999, doi: 10.1016/S0040-4039(99)01780-3.
- [158]K. Balke, M. Kadow, H. Mallin, S. Saß, and U. T. Bornscheuer, “Discovery, application and protein engineering of Baeyer-Villiger monooxygenases for organic synthesis,” *Org. Biomol. Chem.*, vol. 10, no. 31, pp. 6249–6265, 2012, doi: 10.1039/c2ob25704a.
- [159]C. Tolmie, M. S. Smit, and D. J. Opperman, “Native roles of Baeyer-Villiger monooxygenases in the microbial metabolism of natural compounds,” *Nat. Prod. Rep.*, vol. 36, no. 2, pp. 326–353, 2019, doi: 10.1039/c8np00054a.
- [160]I. A. Mirza *et al.*, “Crystal structures of cyclohexanone monooxygenase reveal complex domain movements and a sliding cofactor,” *J. Am. Chem. Soc.*, vol. 131, no. 25, pp. 8848–8854, 2009, doi: 10.1021/ja9010578.
- [161]B. J. Yachnin, T. Sprules, M. B. McEvoy, P. C. K. Lau, and A. M. Berghuis, “The substrate-bound crystal structure of a baeyer-villiger monooxygenase exhibits a criegee-like conformation,” *J. Am. Chem. Soc.*, vol. 134, no. 18, pp. 7788–7795, 2012, doi: 10.1021/ja211876p.
- [162]B. J. Yachnin, M. B. McEvoy, R. J. D. Maccuish, K. L. Morley, P. C. K. Lau, and A. M. Berghuis, “Lactone-bound structures of cyclohexanone monooxygenase provide insight into the stereochemistry of catalysis,” *ACS Chem. Biol.*, vol. 9, no. 12, pp. 2843–2851, 2014, doi: 10.1021/cb500442e.
- [163]D. E. Torres Pazmiño, R. Snajdrova, D. V. Rial, M. D. Mihovilovic, and M. W. Fraaije, “Altering the substrate specificity and enantioselectivity of phenylacetone monooxygenase by structure-inspired enzyme redesign,” *Adv. Synth. Catal.*, vol. 349, no. 8–9, pp. 1361–1368, 2007, doi: 10.1002/adsc.200700045.
- [164]M. Gibson, M. Nur-E-Alam, F. Lipata, M. A. Oliveira, and J. Rohr, “Characterization of kinetics and products of the Baeyer-Villiger oxygenase MtmOIV, the key enzyme of the biosynthetic pathway
-

-
- toward the natural product anticancer drug mithramycin from *Streptomyces argillaceus*,” *J. Am. Chem. Soc.*, vol. 127, no. 50, pp. 17594–17595, 2005, doi: 10.1021/ja055750t.
- [165]M. A. Bosserman, T. Downey, N. Noinaj, S. K. Buchanan, and J. Rohr, “Molecular insight into substrate recognition and catalysis of baeyer-villiger monooxygenase MtmOIV, the key frame-modifying enzyme in the biosynthesis of anticancer agent mithramycin,” *ACS Chem. Biol.*, vol. 8, no. 11, pp. 2466–2477, 2013, doi: 10.1021/cb400399b.
- [166]M. C. Tang, H. Y. He, F. Zhang, and G. L. Tang, “Baeyer-villiger oxidation of acyl carrier protein-tethered thioester to acyl carrier protein-linked thiocarbonate catalyzed by a monooxygenase domain in FR901464 biosynthesis,” *ACS Catal.*, vol. 3, no. 3, pp. 444–447, 2013, doi: 10.1021/cs300819e.
- [167]J. Jiang *et al.*, “Genome mining in *Streptomyces avermitilis*: A biochemical Baeyer-Villiger reaction and discovery of a new branch of the pentalenolactone family tree,” *Biochemistry*, vol. 48, no. 27, pp. 6431–6440, 2009, doi: 10.1021/bi900766w.
- [168]V. Alphand, G. Carrea, R. Wohlgemuth, R. Furstoss, and J. M. Woodley, “Towards large-scale synthetic applications of Baeyer-Villiger monooxygenases,” *Trends Biotechnol.*, vol. 21, no. 7, pp. 318–323, 2003, doi: 10.1016/S0167-7799(03)00144-6.
- [169]M. W. Fraaije, N. M. Kamerbeek, W. J. H. Van Berkel, and D. B. Janssen, “Identification of a Baeyer-Villiger monooxygenase sequence motif,” *FEBS Lett.*, vol. 518, no. 1–3, pp. 43–47, 2002, doi: 10.1016/S0014-5793(02)02623-6.
- [170]J. Rebehmed, V. Alphand, V. De Berardinis, and A. G. De Brevern, “Evolution study of the Baeyer-Villiger monooxygenases enzyme family: Functional importance of the highly conserved residues,” *Biochimie*, vol. 95, no. 7, pp. 1394–1402, 2013, doi: 10.1016/j.biochi.2013.03.005.
- [171]C. Szolkowy, L. D. Eltis, N. C. Bruce, and G. Grogan, “Insights into sequence-activity relationships amongst Baeyer-Villiger monooxygenases as revealed by the intragenomic complement of enzymes from *Rhodococcus jostii* RHA1,” *ChemBioChem*, vol. 10, no. 7, pp. 1208–1217, 2009, doi: 10.1002/cbic.200900011.
- [172]A. Singh, N. Singh Chauhan, H. V. Thulasiram, V. Taneja, and R. Sharma, “Identification of two flavin monooxygenases from an effluent
-

-
- treatment plant sludge metagenomic library,” *Bioresour. Technol.*, vol. 101, no. 21, pp. 8481–8484, 2010, doi: 10.1016/j.biortech.2010.06.025.
- [173] D. Bonsor *et al.*, “Ligation independent cloning (LIC) as a rapid route to families of recombinant biocatalysts from sequenced prokaryotic genomes,” *Org. Biomol. Chem.*, vol. 4, no. 7, pp. 1252–1260, 2006, doi: 10.1039/b517338h.
- [174] M. L. Mascotti, M. Kurina-Sanz, M. Juri Ayub, and M. W. Fraaije, “Insights in the kinetic mechanism of the eukaryotic Baeyer-Villiger monooxygenase BVMOAf1 from *Aspergillus fumigatus* Af293,” *Biochimie*, vol. 107, no. PB, pp. 270–276, 2014, doi: 10.1016/j.biochi.2014.09.005.
- [175] A. Riebel, H. M. Dudek, G. De Gonzalo, P. Stepniak, L. Rychlewski, and M. W. Fraaije, “Expanding the set of rhodococcal Baeyer-Villiger monooxygenases by high-throughput cloning, expression and substrate screening,” *Appl. Microbiol. Biotechnol.*, vol. 95, no. 6, pp. 1479–1489, 2012, doi: 10.1007/s00253-011-3823-0.
- [176] M. L. Mascotti, W. J. Lapadula, and M. J. Ayub, “The origin and evolution of Baeyer - Villiger Monooxygenases (BVMOs): An ancestral family of flavin monooxygenases,” *PLoS One*, vol. 10, no. 7, pp. 1–16, 2015, doi: 10.1371/journal.pone.0132689.
- [177] S. D. Doig *et al.*, “Reactor operation and scale-up of whole cell Baeyer-Villiger catalyzed lactone synthesis,” *Biotechnol. Prog.*, vol. 18, no. 5, pp. 1039–1046, 2002, doi: 10.1021/bp0200954.
- [178] C. V. F. Baldwin, R. Wohlgemuth, and J. M. Woodley, “The first 200-L scale asymmetric Baeyer-Villiger oxidation using a whole-cell biocatalyst,” *Org. Process Res. Dev.*, vol. 12, no. 4, pp. 660–665, 2008, doi: 10.1021/op800046t.
- [179] M. Bučko, P. Gemeiner, A. Schenk Mayerová, T. Krajčovič, F. Rudroff, and M. D. Mihovilovič, “Baeyer-Villiger oxidations: biotechnological approach,” *Appl. Microbiol. Biotechnol.*, vol. 100, no. 15, pp. 6585–6599, 2016, doi: 10.1007/s00253-016-7670-x.
- [180] H. E. M. Law, C. V. F. Baldwin, B. H. Chen, and J. M. Woodley, “Process limitations in a whole-cell catalysed oxidation: Sensitivity analysis,” *Chem. Eng. Sci.*, vol. 61, no. 20, pp. 6646–6652, 2006, doi: 10.1016/j.ces.2006.06.007.
- [181] J. Lima-Ramos, W. Neto, and J. M. Woodley, “Engineering of

-
- biocatalysts and biocatalytic processes,” *Top. Catal.*, vol. 57, no. 5, pp. 301–320, 2014, doi: 10.1007/s11244-013-0185-0.
- [182] S. D. Doig, L. M. O’Sullivan, S. Patel, J. M. Ward, and J. M. Woodley, “Large scale production of cyclohexanone monooxygenase from *Escherichia coli* TOP10 pQR239,” *Enzyme Microb. Technol.*, vol. 28, no. 2–3, pp. 265–274, 2001, doi: 10.1016/S0141-0229(00)00320-3.
- [183] S. A. Meeuwissen, A. Rioz-Martínez, G. De Gonzalo, M. W. Fraaije, V. Gotor, and J. C. M. Van Hest, “Cofactor regeneration in polymersome nanoreactors: Enzymatically catalysed Baeyer-Villiger reactions,” *J. Mater. Chem.*, vol. 21, no. 47, pp. 18923–18926, 2011, doi: 10.1039/c1jm12407b.
- [184] E. Yabuuchi, I. Yano, H. Hotta, Y. Nishiuchi, and Y. Kosako, “Transfer of Two Burkholderia and an Alcaligenes Species to Ralstonia Gen. Nov.: Proposal of Ralstonia pickettii (Ralston, Palleroni and Doudoroff 1973) Comb. Nov., Ralstonia solanacearum (Smith 1896) Comb. Nov. and Ralstonia eutropha (Davis 1969) Comb. No.,” *Microbiol. Immunol.*, vol. 39, no. 11, pp. 897–904, 1995, doi: 10.1111/j.1348-0421.1995.tb03275.x.
- [185] H. G. Schlegel, H. Kaltwasser, and G. Gottschalk, “Ein Submersverfahren zur Kultur wasserstoffoxydierender Bakterien: Wachstumsphysiologische Untersuchungen,” *Arch. Mikrobiol.*, vol. 38, no. 3, pp. 209–222, 1961, doi: 10.1007/BF00422356.
- [186] S. Alagesan, N. P. Minton, and N. Malys, “13 C-assisted metabolic flux analysis to investigate heterotrophic and mixotrophic metabolism in *Cupriavidus necator* H16,” *Metabolomics*, vol. 14, no. 1, pp. 1–10, 2018, doi: 10.1007/s11306-017-1302-z.
- [187] A. Pohlmann *et al.*, “Genome sequence of the bioplastic-producing ‘Knallgas’ bacterium *Ralstonia eutropha* H16,” *Nat. Biotechnol.*, vol. 24, no. 10, pp. 1257–1262, 2006, doi: 10.1038/nbt1244.
- [188] E. Schwartz and B. Friedrich, “A physical map of the megaplasmid pHG1, one of three genomic replicons in *Ralstonia eutropha* H16,” *FEMS Microbiol. Lett.*, vol. 201, no. 2, pp. 213–219, 2001, doi: 10.1016/S0378-1097(01)00265-8.
- [189] E. Schwartz *et al.*, “A proteomic view of the facultatively chemolithoautotrophic lifestyle of *Ralstonia eutropha* H16,” *Proteomics*, vol. 9, no. 22, pp. 5132–5142, 2009, doi: 10.1002/pmic.200900333.
-

-
- [190]R. Cramm, "Genomic view of energy metabolism in *Ralstonia eutropha* H16," *J. Mol. Microbiol. Biotechnol.*, vol. 16, no. 1–2, pp. 38–52, 2008, doi: 10.1159/000142893.
- [191]C. Bi *et al.*, "Development of a broad-host synthetic biology.pdf," pp. 1–10, 2013.
- [192]B. Bowien and B. Kusian, "Genetics and control of CO₂ assimilation in the chemoautotroph *Ralstonia eutropha*," *Arch. Microbiol.*, vol. 178, no. 2, pp. 85–93, 2002, doi: 10.1007/s00203-002-0441-3.
- [193]T. Burgdorf *et al.*, "[NiFe]-hydrogenases of *Ralstonia eutropha* H16: Modular enzymes for oxygen-tolerant biological hydrogen oxidation," *J. Mol. Microbiol. Biotechnol.*, vol. 10, no. 2–4, pp. 181–196, 2006, doi: 10.1159/000091564.
- [194]J. Ratzka, L. Lauterbach, O. Lenz, and M. B. Ansorge-Schumacher, "Systematic evaluation of the dihydrogen-oxidising and NAD⁺-reducing soluble [NiFe]-hydrogenase from *Ralstonia eutropha* H16 as a cofactor regeneration catalyst," *Biocatal. Biotransformation*, vol. 29, no. 6, pp. 246–252, 2011, doi: 10.3109/10242422.2011.615393.
- [195]S. Grunwald *et al.*, "Kinetic and stoichiometric characterization of organoautotrophic growth of *Ralstonia eutropha* on formic acid in fed-batch and continuous cultures," *Microb. Biotechnol.*, vol. 8, no. 1, pp. 155–163, 2015, doi: 10.1111/1751-7915.12149.
- [196]I. Hilker, M. C. Gutiérrez, V. Alphand, R. Wohlgemuth, and R. Furstoss, "Microbiological transformations 57. Facile and efficient resin-based in situ SFPR preparative-scale synthesis of an enantiopure 'unexpected' lactone regioisomer via a baeyer-villiger oxidation process," *Org. Lett.*, vol. 6, no. 12, pp. 1955–1958, 2004, doi: 10.1021/ol049508n.
- [197]S. Lütte *et al.*, "Autotrophic production of stable-isotope-labeled arginine in *Ralstonia eutropha* strain H16," *Appl. Environ. Microbiol.*, vol. 78, no. 22, pp. 7884–7890, 2012, doi: 10.1128/AEM.01972-12.
- [198]G. C. Barnard, G. E. Henderson, S. Srinivasan, and T. U. Gerngross, "High level recombinant protein expression in *Ralstonia eutropha* using T7 RNA polymerase based amplification," *Protein Expr. Purif.*, vol. 38, no. 2, pp. 264–271, 2004, doi: 10.1016/j.pep.2004.09.001.
- [199]J. M. Park, Y. S. Jang, T. Y. Kim, and S. Y. Lee, "Development of a gene knockout system for *Ralstonia eutropha* H16 based on the broad-host-range vector expressing a mobile group II intron," *FEMS Microbiol.*
-

-
- Lett.*, vol. 309, no. 2, pp. 193–200, 2010, doi: 10.1111/j.1574-6968.2010.02041.x.
- [200] S. Gruber *et al.*, “Design of inducible expression vectors for improved protein production in *Ralstonia eutropha* H16 derived host strains,” *J. Biotechnol.*, vol. 235, pp. 92–99, 2016, doi: 10.1016/j.jbiotec.2016.04.026.
- [201] D. Przybylski, T. Rohwerder, C. Dilßner, T. Maskow, H. Harms, and R. H. Müller, “Exploiting mixtures of H₂, CO₂, and O₂ for improved production of methacrylate precursor 2-hydroxyisobutyric acid by engineered *Cupriavidus necator* strains,” *Appl. Microbiol. Biotechnol.*, vol. 99, no. 5, pp. 2131–2145, 2015, doi: 10.1007/s00253-014-6266-6.
- [202] S. E. Nybo, N. E. Khan, B. M. Woolston, and W. R. Curtis, “Metabolic engineering in chemolithoautotrophic hosts for the production of fuels and chemicals,” *Metab. Eng.*, vol. 30, pp. 105–120, 2015, doi: 10.1016/j.ymben.2015.04.008.
- [203] Y. C. Chen, O. P. Peoples, and C. T. Walsh, “Acinetobacter cyclohexanone monooxygenase: gene cloning and sequence determination,” *J. Bacteriol.*, vol. 170, no. 2, pp. 781–789, 1988, doi: 10.1128/jb.170.2.781-789.1988.
- [204] N. M. Kamerbeek, M. J. H. Moonen, J. G. M. Van Der Ven, W. J. H. Van Berkel, M. W. Fraaije, and D. B. Janssen, “4-Hydroxyacetophenone monooxygenase from *Pseudomonas fluorescens* ACB. A novel flavoprotein catalyzing Baeyer-Villiger oxidation of aromatic compounds,” *Eur. J. Biochem.*, vol. 268, no. 9, pp. 2547–2557, 2001, doi: 10.1046/j.1432-1327.2001.02137.x.
- [205] H. Iwaki, Y. Hasegawa, S. Wang, M. M. Kayser, and P. C. K. Lau, “Cloning and characterization of a gene cluster involved in cyclopentanol metabolism in *Comamonas* sp. strain NCIMB 9872 and biotransformations effected by *Escherichia coli*-expressed cyclopentanone 1,2-monooxygenase,” *Appl. Environ. Microbiol.*, vol. 68, no. 11, pp. 5671–5684, 2002, doi: 10.1128/AEM.68.11.5671-5684.2002.
- [206] E. Malito, A. Alfieri, M. W. Fraaije, and A. Mattevi, “Crystal structure of a Baeyer-Villiger monooxygenase,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 101, no. 36, pp. 13157–13162, 2004, doi: 10.1073/pnas.0404538101.
- [207] M. A. Larkin *et al.*, “Clustal W and Clustal X version 2.0,” *Bioinformatics*, vol. 23, no. 21, pp. 2947–2948, 2007, doi:
-

10.1093/bioinformatics/btm404.

- [208]H. Lilie, E. Schwarz, and R. Rudolph, "Advances in refolding of proteins produced in *E. coli*," *Curr. Opin. Biotechnol.*, vol. 9, no. 5, pp. 497–501, 1998, doi: 10.1016/S0958-1669(98)80035-9.
- [209]N. Ferrer-Miralles, J. Domingo-Espín, J. Corchero, E. Vázquez, and A. Villaverde, "Microbial factories for recombinant pharmaceuticals," *Microb. Cell Fact.*, vol. 8, pp. 1–8, 2009, doi: 10.1186/1475-2859-8-17.
- [210]D. H. Lee, M. D. Kim, W. H. Lee, J. H. Seo, and D. H. Kweon, "Consortium of fold-catalyzing proteins increases soluble expression of cyclohexanone monooxygenase in recombinant *Escherichia coli*," *Appl. Microbiol. Biotechnol.*, vol. 63, no. 5, pp. 549–552, 2004, doi: 10.1007/s00253-003-1370-z.
- [211]D. Baeyer-villiger, K. H. Jones, R. O. Y. T. Smith, and W. Trudgill, "from camphor-grown *Pseudomonas putida* ATCC 17453," pp. 797–805, 1993.
- [212]K. Ikura *et al.*, "Co-overexpression of folding modulators improves the solubility of the recombinant guinea pig liver transglutaminase expressed in *Escherichia coli*," *Prep. Biochem. Biotechnol.*, vol. 32, no. 2, pp. 189–205, 2002, doi: 10.1081/PB-120004130.
- [213]K. Nishihara, M. Kanemori, M. Kitagawa, H. Yanagi, and T. Yura, "Chaperone coexpression plasmids: Differential and synergistic roles of DnaK-DnaJ-GrpE and GroEL-GroES in assisting folding of an allergen of Japanese cedar pollen, Cryj2, in *Escherichia coli*," *Appl. Environ. Microbiol.*, vol. 64, no. 5, pp. 1694–1699, 1998, doi: 10.1128/aem.64.5.1694-1699.1998.
- [214]J. Klose *et al.*, "Hexa-histidin tag position influences disulfide structure but not binding behavior of in vitro folded N-terminal domain of rat corticotropin-releasing factor receptor type 2a," *Protein Sci.*, vol. 13, no. 9, pp. 2470–2475, 2004, doi: 10.1110/ps.04835904.
- [215]P. Ledent *et al.*, "Unexpected influence of a C-terminal-fused His-tag on the processing of an enzyme and on the kinetic and folding parameters," *FEBS Lett.*, vol. 413, no. 2, pp. 194–196, 1997, doi: 10.1016/S0014-5793(97)00908-3.
- [216]J. W. Jarvik and C. A. Telmer, "Epitope tagging," *Annu. Rev. Genet.*, vol. 32, pp. 601–618, 1998, doi: 10.1146/annurev.genet.32.1.601.

- [217]R. C. Stevens, "Design of high-throughput methods of protein production for structural biology," *Structure*, vol. 8, no. 9, pp. 177–185, 2000, doi: 10.1016/S0969-2126(00)00193-3.
- [218]K. Terpe, "Overview of tag protein fusions: From molecular and biochemical fundamentals to commercial systems," *Appl. Microbiol. Biotechnol.*, vol. 60, no. 5, pp. 523–533, 2003, doi: 10.1007/s00253-002-1158-6.
- [219]D. S. Waugh, "Making the most of affinity tags," *Trends Biotechnol.*, vol. 23, no. 6, pp. 316–320, 2005, doi: 10.1016/j.tibtech.2005.03.012.
- [220]J. Arnau, C. Lauritzen, G. E. Petersen, and J. Pedersen, "Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins," *Protein Expr. Purif.*, vol. 48, no. 1, pp. 1–13, 2006, doi: 10.1016/j.pep.2005.12.002.
- [221]S. M. Smith, *Strategies for the purification of membrane proteins*, vol. 1485. 2017.
- [222]S. Costa, A. Almeida, A. Castro, and L. Domingues, "Fusion tags for protein solubility, purification, and immunogenicity in *Escherichia coli*: The novel Fh8 system," *Front. Microbiol.*, vol. 5, no. FEB, pp. 1–20, 2014, doi: 10.3389/fmicb.2014.00063.
- [223]S. E. Andrew *et al.*, "• © 199 3," vol. 4, no. august, pp. 398–403, 1993.
- [224]E. R. Lavallie, Z. Lu, E. A. Diblasio-Smith, L. A. Collins-Racie, and J. M. Mccoy, "Thioredoxin as a fusion partner for production of soluble recombinant proteins in *Escherichia coli*," *Methods Enzymol.*, vol. 326, no. 1989, pp. 322–340, 2000, doi: 10.1016/s0076-6879(00)26063-1.
- [225]S. K. Katti, D. M. LeMaster, and H. Eklund, "Crystal structure of thioredoxin from *Escherichia coli* at 1.68 Å resolution," *J. Mol. Biol.*, vol. 212, no. 1, pp. 167–184, 1990, doi: 10.1016/0022-2836(90)90313-B.
- [226]A. Kirschner, J. Altenbuchner, and U. T. Bornscheuer, "Cloning, expression, and characterization of a Baeyer-Villiger monooxygenase from *Pseudomonas fluorescens* DSM 50106 in *E. coli*," *Appl. Microbiol. Biotechnol.*, vol. 73, no. 5, pp. 1065–1072, 2007, doi: 10.1007/s00253-006-0556-6.
- [227]J. Rehdorf, A. Kirschner, and U. T. Bornscheuer, "Cloning, expression and characterization of a Baeyer-Villiger monooxygenase from *Pseudomonas putida* KT2440," *Biotechnol. Lett.*, vol. 29, no. 9, pp.

1393–1398, 2007, doi: 10.1007/s10529-007-9401-y.

- [228]Y. Yoshikane, N. Yokochi, K. Ohnishi, and T. Yagi, “Coenzyme precursor-assisted cooperative overexpression of an active pyridoxine 4-oxidase from *Microbacterium luteolum*,” *Protein Expr. Purif.*, vol. 34, no. 2, pp. 243–248, 2004, doi: 10.1016/j.pep.2003.11.013.
- [229]L. Wang and W. Wang, “Coenzyme precursor-assisted expression of a cholesterol oxidase from *Brevibacterium* sp. in *Escherichia coli*,” *Biotechnol. Lett.*, vol. 29, no. 5, pp. 761–766, 2007, doi: 10.1007/s10529-006-9295-0.
- [230]J. G. Thomas and F. Baneyx, “Protein misfolding and inclusion body formation in recombinant *Escherichia coli* cells overexpressing heat-shock proteins,” *J. Biol. Chem.*, vol. 271, no. 19, pp. 11141–11147, 1996, doi: 10.1074/jbc.271.19.11141.
- [231]R. K. Scopes, “Enzyme Activity and Assays,” *Encycl. Life Sci.*, pp. 1–6, 2002, doi: 10.1038/npg.els.0000712.
- [232]A. Ö. Triantafyllou, E. Wehtje, P. Adlercreutz, and B. Mattiasson, “How do additives affect enzyme activity and stability in nonaqueous media?,” *Biotechnol. Bioeng.*, vol. 54, no. 1, pp. 67–76, 1997, doi: 10.1002/(SICI)1097-0290(19970405)54:1<67::AID-BIT8>3.0.CO;2-W.
- [233]K. Talley and E. Alexov, “On the pH-optimum of activity and stability of proteins,” *Proteins Struct. Funct. Bioinforma.*, vol. 78, no. 12, pp. 2699–2706, 2010, doi: 10.1002/prot.22786.
- [234]J. Müller *et al.*, “Engineering of *Ralstonia eutropha* H16 for autotrophic and heterotrophic production of methyl ketones,” *Appl. Environ. Microbiol.*, vol. 79, no. 14, pp. 4433–4439, 2013, doi: 10.1128/AEM.00973-13.

Appendix A Sequence alignment of BVMOs



Figure 64. Multiple nucleotide sequence alignment of CPMO1 from *C.necator* H16 with known BVMOs.



Figure 65. Multiple nucleotide sequence alignment of CPMO2 from *C.necator* H16 with known BVMOs



Figure 66. Multiple nucleotide sequence alignment of CPMO3 from *C.necator* H16 with known BVMOs.



Figure 67. Multiple nucleotide sequence alignment of CPMO4 from *C.necator* H16 with known BVMOs.



Figure 68. Multiple nucleotide sequence alignment of CPMO5 from *C.necator* H16 with known BVMOs.



Figure 69. Multiple nucleotide sequence alignment of CPMO6 from *C.necator* H16 with known BVMOs



Figure 70. Multiple nucleotide sequence alignment of CPMO7 from *C.necator* H16 with known BVMOs.



Figure 71. Multiple nucleotide sequence alignment of CPMO8 from *C.necator* H16 with known BVMOs.



Figure 72. Multiple nucleotide sequence alignment of CPMO9 from *C.necator* H16 with known BVMOs.



Figure 73. Multiple nucleotide sequence alignment of CPMO10 from *C.necator* H16 with known BVMOs.

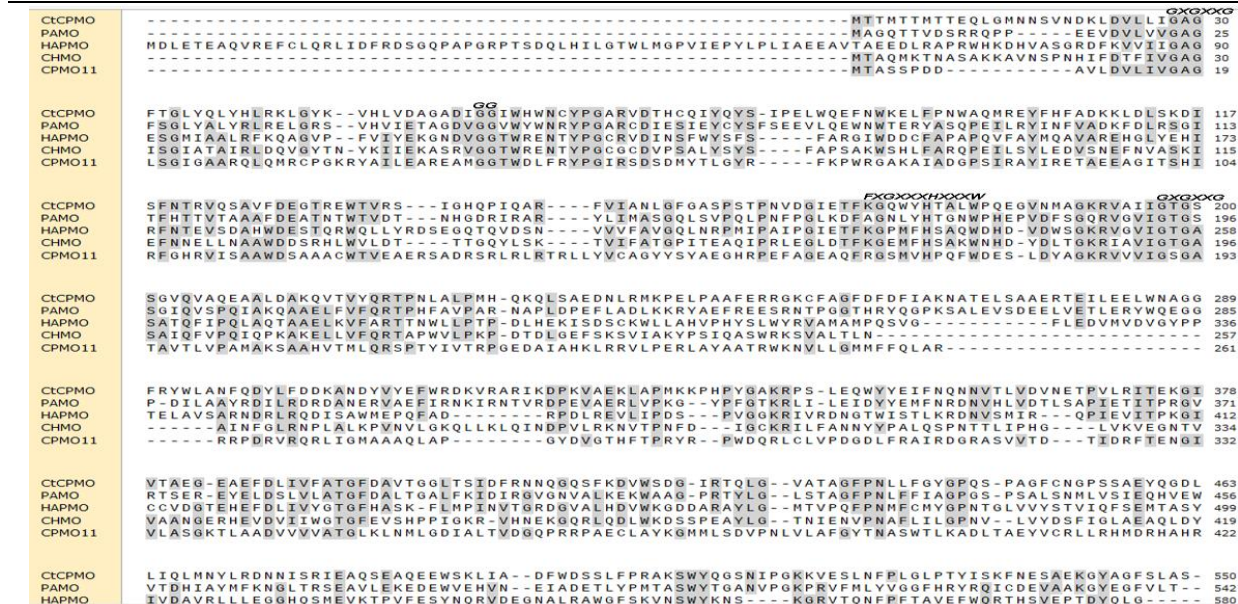


Figure 74. Multiple nucleotide sequence alignment of CPMO11 from *C.necator* H16 with known BVMOs.

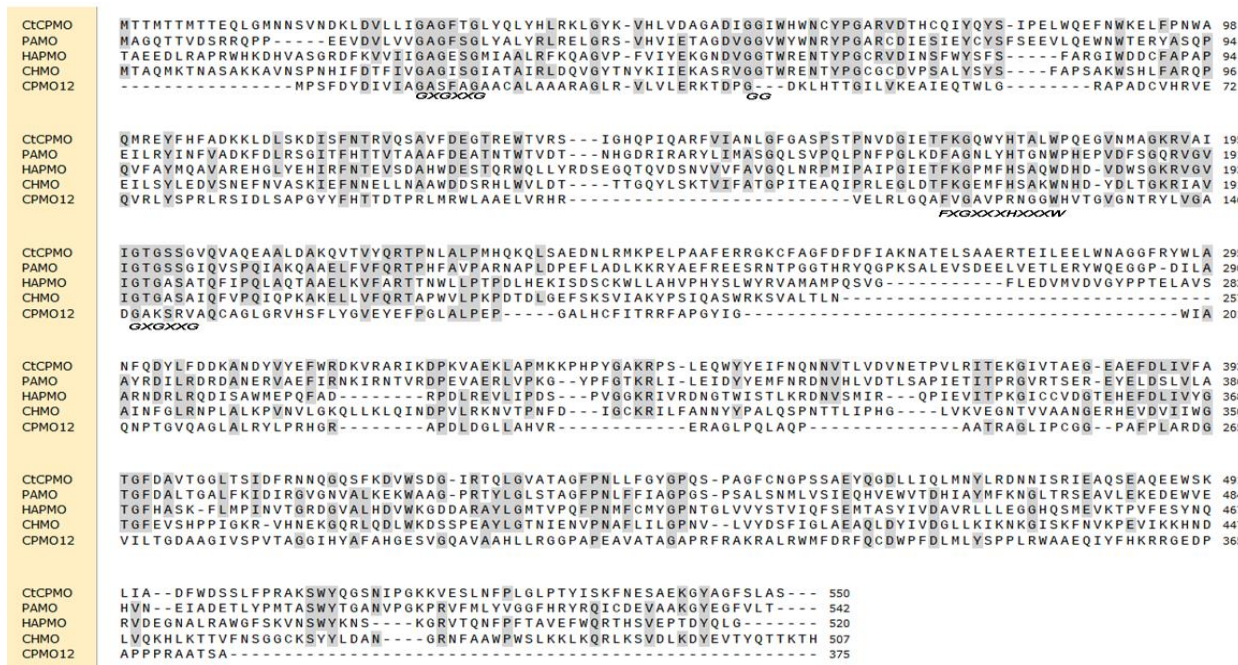


Figure 75. Multiple nucleotide sequence alignment of CPMO12 from *C.necator* H16 with known BVMOs.

CtCPMO	MTTMTTTEQLGMNNSVNDKLDVLLIGAGFTGLYLQYLHLRKLGYK-VHLVDAGADIGGIWHWNCYPGARVDTHCQIYQYS-IPELWQEFNWKELFPNWA	98
PAMO	MAGQTTVDSRRQP-----EEVDLVVVGAGFSGLYALYRLRELGRS-VHVIETAGDVGGVWYNNRYPGARCDIESIEYCSFSEEVLEQWNTERYASQP	94
HAPMO	TAEDLRAPRWKHDHVASGRDFKVVIIIGAGESGMIAALRFKQAGVP-FVYIEKGNVDVGGTWRENTYPGCRVDINSFWYSF-----FARGIWDCCFAPAP	94
CHMO	MTAQMKTNASAKKAVNSPNHIFDFTFVVGAGISGIATAIRLDQVGYTVNIIIEKASRVGGTWRENTYPGCGCDVPSALYSYS-----FAPSAKWSHLFARQP	96
CPMO13	-QAQFVRNQPVLPRGRPTERIELAIIIGAGMSGIAAAIQAKDRGFR-VRIEFSNNKVGWVAANDYPGVAVDTPATVYYSLS-----VELNPSWSNYPVGS	94
	GG	
CtCPMO	QMREYFFHADKKLDSLKDISFNTRVQSAVDFEGTREWTVRS----IGHQPIQARFVIANLGFGASPSTPNVDGIETFKGQWYHTALWPQEGVNMAGKRVA	194
PAMO	EILRYINFVADKFDLRSGITFHTTAAAFDEATNTWTDV----NHGDRIRARYLIMASGQLSVPQLPNFPGLKDFAGNLYHTGNWPHEPVDVFSGQRVQ	190
HAPMO	QVFAYMQAVAREHGLYEHIRFNTVSDAHWDESTQRWQLLYRD-SEGQTQVDSNVVVFVAVGQLNRMPAIPGIIETFKGPMFHSQAQWHDH-VDWSGKRQV	192
CHMO	EILSYLEDVSNFVASKIEFNELLNAAWDDSRHLVWLD-----TTGQYLSKTVIFATGPIEAQIPRLEGLDTFKGEMFHSKWNHD-YDLTGKRIA	190
CPMO13	EYLRYLEGIVEKHNISEFIELESEILKIQWIEEDQEWELMVVKKGREASRVRAVMSCLGHLNRNPYDLQGRETFKGVSIHANRWKH-VDLRGRKRV	193
	FXGXXXHXOXW	
CtCPMO	IIGTGSSGVQVAEALDAKQVTVYQRTPNLALPMHQKQLSAEDNLRMKPELPAAFERRGKCFAGDFDFDIA-KNATELSAAERTEILEELWNAAGFRYV	293
PAMO	VIGTGSSGIQVSPQIAKQAAELVFVQRTPHFAVPARNAPLDPEFLADLKKRYAEFREESRNPGGTHRYQGP-KSALEVSDEELVETLERYWQEGGP-DI	288
HAPMO	VIGTGASATQIPQLAQTAELKVFARTTNWLLPTPDLHEKISDSCKWLLAHVPHYSLWYRVAMAMPQSVGF-LEDVMDVGVYPPTELAVSARNDRRLRQD	291
CHMO	VIGTGASAIQFVPIQPKAKELLVFQRTAPWVLPKPDTLG-EFSKSVIAKYPSIQASWRKSVALTNLAINF-GLRNPLAKPVNVLGKLLKQLIN---	285
CPMO13	IIGTGATGVQVIGKISQEVHDHVTVMRQPMWIIPIQAGGDIPIQSKRWARQVLPYFLHWRDLKSYWTVSDKVGYPVIRADPEWAKTHVSIISPANDRLMQF	293
	GXGXXG	
CtCPMO	LAFNQDYLDDKANDVYVEFRDKVARIKDPKVAEKLAPMKKPHPYGAKRPSLEQWYIEFNQNNVTLVDVNETPVLRITEKGIPTAE-GEAEFDLIVF	392
PAMO	LAAYRDILRDRDANERVAEFIRNKIRNTVRDPEVAERLVKPG--YFPGTKRLILEIDYEMFNDRNVHLVDTLSAPIETITPRGVRTSE-REVELDSLVL	385
HAPMO	ISAWMEPQFAD-----RPDLREVLIPDS-----PVGGKRVNRDNGTWISTLKR--DNVSMIRQPIEVITPKGICCVGTEHEFDLIVY	367
CHMO	-----DPVLRKNVTPNFD-----IGCKRILFANNYPALQSP--NTLLPHGLVKVEGNTVVAANGRHEVDVVIW	349
CPMO13	CINYINSCFGE-----GSELARKLTPDYAPLGRKSRVDPYDFAPGGFYALSQPNVALETSKLARVVPDGLITVDGKLIELDVIY	374
CtCPMO	ATGFDVAVTGGTSDIFRNNQGSFKDVWSDG-IRTQLGVATAGFPNLLFGYGPQS-PAGFCNGPSSAEYQGDLLIQLMNYLRDNNISRIEAEQSEAEQW	490
PAMO	ATGFDALGALFKIDIRGVGNVALKEKWAAG-PRTYLGLSTAGFPNLLFIAGPQS-PSALSNMNVLSIEQHVWVTDHIAVMFKNGLTRSEAWLEKEDWV	483
HAPMO	GTGFHASK-FLMPINVTGRDGVALHDVWGGDARAYLGHTVPOFPMFCMYGPNITGLVYVSTVQFSEMTASVIVDAVRLLEGGHQSMEVKTTPVFESYN	466
CHMO	GTGFVESHPPIGKR-VHNEKQRLQDLWKDSSPEAYLGTNIENVPNAFLILGNV--LVYDSFGLAEALDYVDGLLKIKNKGISKFNKPEVVKKH	446
CPMO13	ATG--LTDWLSPTEVAGRNGVRLSDVWRDNNPSSYLGGMVKFPNLFINSQPTGAGHAGGHMFAEVVNHAMECLQLLVEEDARSLEVTQEAFDEHN	472
CtCPMO	KLIA--DFWDSFLFRAKSWYQGSNIPGKVESLNFPLGLPTYISKFNESAEGYAGFSLAS---550	
PAMO	ERVN--EIADETLYPMTASWYTGANVPGKPRVFMVYVGGFHYRQICDEVAAGYEGFVLT---542	
HAPMO	QRVDEGNALRAWGFSKVNWSYKNS---KGRVTQNFPTAVEFVQRTHSVEPTDYLQ---520	
CHMO	DLVQKHLKTTVFNSSGGCKSYLLDAN---GRNFAAWPWSLKKLQRLKSVLDKDYEVTVYTTKTH	507
CPMO13	RKLDEKMAGAIWAWERRAHTYTNQ---KGRPILPTWRHVDVFRMTRPEMKEAIFLR-----527	

Figure 76. Multiple nucleotide sequence alignment of CPMO13 from *C.necator* H16 with known BVMOs.

CtCPMO	-----MTTMTTTEQLGMNNSVNDKLDVLLIGAGFTGLYLQYLHL	40
PAMO	-----MAGQTTVDSRRQP-----EEVDLVVVGAGFSGLYALYRL	35
HAPMO	HDLETEAQVREFCLQLRIDFRDSGGQAPGRPTSQDLHILGTWLNQVPIEYPLPLTAAEAVTAEDLRAPRWKHDHVASGRDFKVVIIIGAGESGMIAALRF	100
CHMO	-----MTAQMKTNASAKKAVNSPNHIFDFTFVVGAGISGIATAIRL	40
CPMO14	-----MKTETTDVVIIIGAGPAGSVAAGLL	24
	GXGXXG	
CtCPMO	RKLGYK-VHLVDAGADIGGIWHWNCYPGARVDTHCQIYQYS-IPELWQEFNWKELFPNWAQMREYFFHADKKLDSLKDISFNTRVQSAVDFEGTREWTVR	138
PAMO	RELGRS-VHVIETAGDVGGVWYNNRYPGARCDIESIEYCSFSEEVLEQWNTERYASQPEILRYINFVADKFDLRSGITFHTTAAAFDEATNTWTDV	134
HAPMO	KQAGVP-FVYIEKGNVDVGGTWRENTYPGCRVDINSFWYSF-----FARGIWDCCFAPAPQVFAYMQAVAREHGLYEHIRFNTVSDAHWDESTQRWQL	194
CHMO	DQVGYTNYKIIIEKASRVGGTWRENTYPGCGCDVPSALYSYS-----FAPSAKWSHLFARQPEILSYLEDVSNFVASKIEFNELLNAAWDDSRHLVWLD	136
CPMO14	RKHGIP-VLVIEKETFPFRSFIGESLLPQSMAYIEAG-----MLRAVVEAGFYKNGAFAARGDQQTAFDFREKFSFGWG---TTYVQV	104
	GG	
CtCPMO	S---IGHQPIQARFVIANLGFGASPSTPNVDGIETFKGQWYHTALWPQEGVNMAGKRVAIIGTGSSGVQVAEALDAKQVTVYQRTPNLALPMHQKQLS	235
PAMO	T---NHGDRIRARYLIMASGQLSVPQLPNFPGLKDFAGNLYHTGNWPHEPVDVFSGQRVGVIGTGSSGIQVSPQIAKQAAELVFVQRTPHFAVPARNAPLD	231
HAPMO	YRDSEGGTQVDSNVVVFVAVGQLNRMPAIPGIIETFKGPMFHSQAQWHDH-VDWSGKRQVGVIGTGASATQIPQLAQTAELKVFARTTNWLLPTPDLHEK	293
CHMO	T-----TTGQYLSKTVIFATGPIEAQIPRLEGLDTFKGEMFHSKWNHD-YDLTGKRIAIVIGTGASAIQFVPIQPKAKELLVFQRTAPWVLPKPDTLG	231
CPMO14	R-----ARFDDVLIIEAARQGAEVRFSLVEDVDVSSGAQPEVTVRAPDGRYS-VRAKFLLDASGFGRLRRLQLLESFSGFVPSALFTHVEDRIAPG	197
	FXGXXXHXOXW	
CtCPMO	AEDNLRMKPELPAAFERRGKCFAGDFDFDIAKNAATELSAAERTEILEELWNAAGFRYWLAFNQDYLDDKANDVYVEFRDKVARIKDPKVAEKLAPMK	335
PAMO	PEFLADLKKRYAEFREESRNPGGTHRYQGP-KSALEVSDEELVETLERYWQEGGP-DILAAYRDILRDRDANERVAEFIRNKIRNTVRDPEVAERLVKPG	330
HAPMO	ISDSCKWLLAHVPHYSLWYRVAMAMPQSVG-----FLEDVMDVGVYPPTELAVSARNDRRLQDISAWMEPQFAD-----RPDLREVLIPDS	374
CHMO	EFSKSVIAKYPSIQASWRKSVALTNLAINF-GLRNPLAKPVNVLGKLLKQLINDVLRKNVTPNF	297
CPMO14	AFDRNKILISVHPQH-----AINFGLRNPLAKPVNVLGKLLKQLINDVLRKNVTPNF	223
CtCPMO	KHPYGAKRPS-LEQWYIEFNQNNVTLVDVNETPVLRITEKGIPTAE-GEAEFDLIVFATGFDVAVTGGTSDIFRNNQGSFKDVWSDG-IRTQLGVAT	432
PAMO	--YFPGTKRLI-LEIDYEMFNDRNVHLVDTLSAPIETITPRGVRTSER-EYELDSLVLATGFDALGALFKIDIRGVGNVALKEKWAAG-PRTYLGLST	425
HAPMO	D---PVGGKRVNRDNGTWISTLKRDNVSHIR---QPIEVITPKGICCVGTEHEFDLIVYGTGFHASK-FLMPINVTGRDGVALHDVWGGDARAYLGHTV	467
CHMO	D-----IGCKRILFANNYPALQSPNTTLPHG---LVKVEGNTVVAANGRHEVDVVIWIGTGFVESHPPIGKR-VHNEKQRLQDLWKDSSPEAYLGTNI	389
CPMO14	D---GRCSSGVVAEKAFDRYGTETELR---ALVAEPELAKLLGNAQWDTPARQIAQYSANVKSLEWNGYALLGN-----AGEFLD	301
CtCPMO	AGFPNLLFGYGPQS-PAGFCNGPSSAEYQGDLLIQLMNYLRDNNISRIEAEQSEAEQW	529
PAMO	AGFPNLEFFIAGPQS-PSALSNMNVLSIEQHVWVTDHIAVMFKNGLTRSEAWLEKEDWV	522
HAPMO	PQFPMFCMYGPNITGLVYVSTVQFSEMTASVIVDAVRLLEGGHQSMEVKTTPVFESYNQRVDEGNALRAWGFSKVNWSYKNS---KGRVTQNFPTAVE	563
CHMO	ENVPNAFLILGNV--LVYDSFGLAEALDYVDGLLKIKNKGISKFNKPEVVKHNDLVQKHLKTTVFNSSGGCKSYLLDAN---GRNFAAWPWSL	483
CPMO14	PVESSGVTIAFKSASLAAQCLVRQLAGDVBQEQFAQPKAGVDCFRVFEAWYEGRFQQLIFHPNANSEIRDMISSLQAGYAWDRSNPFAVEPRRRLA	401
CtCPMO	TYISKFNESAEGYAGFSLAS---550	
PAMO	RYRQICDEVAAGYEGFVLT---542	
HAPMO	<	

Figure 77. Multiple nucleotide sequence alignment of CPMO14 from *C.necator* H16 with known BVMOs.



Figure 78. Multiple nucleotide sequence alignment of CPMO15 from *C.necator* H16 with known BVMOs.



Figure 79. Multiple nucleotide sequence alignment of CPMO16 from *C.necator* H16 with known BVMOs.



Figure 80. Multiple nucleotide sequence alignment of CPMO17 from *C.necator* H16 with known BVMOs.

Table 41. The amino acid sequence used in this study

CPMO0
MNAPAE L P Q Q A S P A L E H T D I A I I G S G F A G L G M A I R L R E R G E R N F L I F E K A G S V G G T W R D N H Y P G C A C D V Q S H L Y S F S F A P N P D W S R M F S T Q P E I R A Y L E G C T D R F G L R P H L R L H H E L R H A A W D E A A Q V W R L E M A D G K V V R A R V L V S G M G G L S R P S Y P N I P G L D S F E G E C F H S Q H W N H A Y A L A G K R V A V I G T G A S A I Q F V P Q I A P Q V A Q L D L Y Q R T P P W I M P K P D R E V S A T E R W L F R R L P F T Q T L V R T G I Y W M L E S R V L G F V L H P K I M K V V E R V A R R H L A R Q V R D P A L R A T L T P D Y T I G C K R V L I S N D Y Y P A L T R D N V D V T T G I A R V E R D A V V L R D G T R R P A D C I I L G T G F H A T D P L P R G V I S G L D G V D L V D A W R D G A Q A Y L G T T V A G F P N L F F V V G P N T G L G H S S M V F M I E S Q V A Y I A D A L Q Q M R R E G V Q A V D V R E P V Q R R F N E R L Q A R L G H A I W S A G G C A S W Y I D P R S G R N T T L W P G F T W Q F R R A T A S F R M T D Y D L Y P V R R A L P D E T A M Q P G P A D I A T V A G A
CPMO1
M S T T L T P T T V A F I G L G A M G S H M V R H L L A A G H T V R A F V R R P E A A E A A R A L G A E P F F T P A E A A R G A S V V F T N V T S S E D V R E V L L G E Q G V I H G A A P G T I C V D H S T I S P I V T R E I A A A L A A R G I E A L D C P V S G G T M G A E A G T L T I M V G G K P E M L E R V R P L L Q Q L G K T I T H I G D H G A G Q V A K L C N Q I A Q V V N I E G I A E A M R F A A A Q N V D T G R V F E A M A T G M A G S R M L D L M G P K M V A R N F A A G I E A R L H D K D F G L A R D I A E E I G L D L P A M Q A T S A Q L R T L M A N G W G K D D T S S L L R V L E G
CPMO2
M S N Y D V V V G A G F T G L T A A Y A L A K Q G K K V H V V E A D A T P G G L A G T F E F A D G V K L E K F Y H H W F N N D V Y V P E L V K E L G M E G D V I L L P T R T G M Y F N G R M W K L S T P L D L L K F K A L S F I D R I R L G L L V F Q V R R I K D W K T I E H L S I R E W L E P I C G K T V Y R I V W E P L I S S K F S V Y A E A V N A V W M W K K L V L R G S T R N D K G G E E L A Y F K G G F G R L A E A M V A A I E R A G G E V T F K N K V N G V L A D G K K I T G L E T E Q G E L K G K Q Y L F T P S F P I I A D M F E G T A P P E W L A R L R R V R Y L G N V C L V L R L N Q S L S E T Y W L N V N D P G F P F V G V I E H T N F D K P E N Y Q G S H I A Y L S R Y L A V E D P V W G Y T D E Q Y V D F A F E H L K R M F P K M D R S W I V E H R V W R S M Y A Q P V T E R Q Y S E Y P S E E T P Y E N G W I S T M A Q I Y P E D R G T N Y A I R E G K R I V E K L A
CPMO3
M T E T R T P R I V V V G G G A G G L E L V T K L G D K L G R K G Q A Q V V L V D R L P T H I W K P L L H E V A A G S M D P N T H Q L E Y A A Q A R W H H F E F Q L G E L T G I D R I R K T I S V S A S F D E D G A E L L P A R E L P Y D T L V L A I G C V T H F F G V P G A A E N A I A L D V V A Q A E R F R K R L I A A C V R A Q N W R G R V G A D G R P R V D V A I I G A G A T G V E L S A E L R N T A H V L S A Y G L H Q L D P R R D V H I H V I E A G P R I L P A L S E R V S V E T A K L L K K L D V D V L T S E R V T E V T P Q A V L T A S G K H I D A D L T V W A A G I T A P P V L A T L G L P V S R Q G Q I V G P T L Q S E G D P D I F A F G D C A S C P W A E K Q T T V P P R A Q A A H Q Q A T F L Y S A L R A R L D G R P L P S F A F K D L G S L V S L G H F S A V G S L M G G L I G G S M F I E G L M A R V M Y T S L Y R M H V M A L H G A V G M A L D T V S H W L R R K T S P R V K L H
CPMO4
M S G A T S A R T G R Y D V A V L G A G A A G M M C A A V A G Q N G A R V V L I D H A T K L A E K I R I S G G G R C N F T N L Q A G P S N Y L S S N P H F C R S A L A R Y T P Q D F L G L M R R Y G I D W H E K H R G Q L F C N D S A E D V I A M L R A E C D T G H V R W H T G C S V A E V R R E G D D F L L L T A T G M V R A G A L V V A T

GGLSIPKIGATDFGYRIARQFGLGIVETRPALVPLTFDGDWVVPFAPLAGVSLEVDIAT
 GNGKAAGAFREDLLWTHRGLSGPAVLQISSYWRPGTPIAIDLFDGEDAAAWLLEQKT
 GSRKHLGNLLAQRLPARLADAWCAASGVDAAMPLHDVTDKALRKLGLATLNGWRIVP
 SGTEGYRKAEVTLGGVDTRALSSTTMMAREVPGLYFIGEVVDVTGWLGGYNFQWA
 WASAVAAGQAAAEASRTASAGQSVAAQ

CPMO5

MISSTDTVIIGAGPYGLSVSAYLTSAGVPHQILGKPMHAWRNFMPGMLLRSEGFAS
 NLYAPREGYTIEEYCRNRRIEYKPIGMHLPLELVEYGLWFQSQLVSHVRAVEVDM
 HRKDGFFRLALGDGSSLAARRVVMALGLKGFSSQMPALQGMPTPYVVSSEFGSL
 EWAKNKNIVVVGGGQSAIGLAALLSEIGARVRLLRGATVNWSNEPRTSRGGISRL
 RSHVGIGRGWQSNTPSALLSSFLVSEYPIVFHMMNLERRKRTLETGWGPSGAWWL
 RNRVDGKIDILTQTEIRQAEVKNGQVVLQVATRNETSCLSADHVVVATGFKVDMTRH
 SFISQELLNSLSLIDGSPSLTRNFETNVRDLYVVGPAALSFGPALRFIYGAKYAAPQL
 ARHISRCFKNETRISVAQSQUALPSELVEGNNGSGLHGQSILQAPGNME

CPMO6

MLAENNIVIIGGGFAGTTLAQALEKLLPTTHRLILVSDSYTTFNPMMLAEVVGASVFPE
 HVIVPIRQMLRRTQFIMGTVDVNYAQGKLSCTTLAGLREIPFDHLVFAFGTRANLDL
 VPGMAEHALPLKLVGDALFIRNRVLRQLARMELESDPVLRRRLGHFIVVGGGFSGVE
 VAGELADYLCSIRRFYPRVQREELAVTLLHDTGQLLSELPASLGMAAARHMAARGIR
 LRLGTRALRFEADTVVLADGEMIEGATVVCTVGTRPNPLVATLERNLALATRRGRIET
 NPDMSVRGMPGLWAIGDCARVANAATDEDSLPTAQFAVAQARQLAHNLAGHLEQR
 ESKPFSYQSKGAMASIGHMKGVAQVYGFSSGLPAWLAWRALYLMRMPTLGRKLR
 LWAEWTLSMFFPTDITHLRFTRTSEADESPPDGLPPHRPQAPGASATRIATHDFSGI
 PRTPYATPAAPTTLTGQHA

CPMO7

MTITRRDFLNGTALTIAAGLAPVQLLHAQQGGAGVAYPPALTGLRGNHAGTYTLAHS
 LAREGAKYPVVLAREAFDLVVVGGGLSGLAAAWFYRRRFGANRRILIDNHDDFGG
 HAKRNEFTVRGKRLVITYGSAEMPAGAASDAAVRELLAGIGLDATRPPGTVPADPY
 GALGMGRAVFFDAEHFGRDQWVAGDPFGELIDVRSGQQAGAPDAAALSRLDNT
 LPVADRAALARLAAGTTDYLPAISGTERAAALRSMRYASFLRDKAGIGLAGQRFLRS
 RSNDAYALDVDGISVADAIAIGLPAGANMPAPAGNVRLGKSPASRLWFPDGNASLAR
 LLVQSLIPGVASIARPADVVSAAFDYSRLDRDGQPVRLRLNSTAIAVEPDGNITRVTY
 GFSGNLHRVEARHVVLVLAGYNMIPYLLPSLPAKQKEVLHNEAKAPLVYTKVALDHW
 QAFAALRTRRIHAPAMAYTDLWLEPPAGRQPSDPAVLHMLYVPTVPDSGMQARDRF
 RAGRALLGTPFDTLERDIRAQLDRMLGHKGFASKDVIRGITVNRWAHGYSYMPDSL
 AAENVAPDAVWPGLAGVGNISIANSDNAGSPSLTAAVGGQKRAVERLPG

CPMO8

MTDLSAPALTPAAGLAALRLRQDLSWLDLPAKAWTMPRSHDGGQDVLDAVIGGG
 MAGLAACATLTHLGIRVRAFDRAPAGFEGPWATTARMETLRSPKQLTGALGLPALT
 FRAWFEAQFGAAWQTLKIPRLQWMDYLRWYRQVLALDIRNQHQVDAIVPRADGL
 VALSMQTPDGPHTVLARRVVLATGRDGLGGAYVPPVADTLPRHLWAHSSDEMDYA
 KLRGKRVGVVAGASGMDSAATALEGAASVEMLIRRPDIPRVNKSAGNPLTF

GHYGLPDAWKWRLRHYINTQQVPPPRGSTLRVSRHANAYFNLGAALEAIEAQGDVL
 HVQTPQGTFELDFLIFSTGFRIALETRPEFAAFAPHIRFWRDRHAPPAGQEERELSDS
 PDLGAAFEFQEKTPGACPGVTRIHCFYCYPAAALSHGTVSGDIPAISDGARRLGQGIAAL
 LYREDVELHYAALQAYAEPEVFGDEWVPAPPPALRAKA

CPMO9

METVDCWIGAGVVGLAVARALALQGREVIILETENAFGTITSARNSEVIHAGIYYYPAG
 SLKAQLCVRGKAMLYDYCASHHVAHQRCGKLVATSEAQVATLEGIRAKAAANGVD
 DLRLVSRAEAQALEPQLQCHAALLSPSTGIVDSHGLMTALLGDAENAGAMLAVQSPV
 LGGAVTADGIRLEIGSEDGSATTLLARTVVNSAGLTAPELARRIDGMPETHIPPQYYA
 KGCYFTLAGRAPFSRLIYPVPEAAGLGVHLTIDLGGQARFGPNVRWIDEIEYGVDPAD
 ADGFYDEVRRYWPGLADGALQPGYAGIRPKISGPHEVAADFRIDGPAVHGVPGLVH
 LFGIESPGLTSSLIAAERVCAALA

CPMO10

MSTATPLAPTPVTTTDLVIVGAGPVGLFAAFQAGVLGLRCELIDVLDRAAGGQCTELYP
 EKPIYDIPAVPGCLAQDLVDRILLEQCAPFAFPMHFGQRAESLSEIPRQPAPDGHAAH
 ERLLVTTDGGKRFDVSAVLICAGAGAFAPQRVSLPEAPALEGRHVHYAVRDVSRFA
 GKRVVVAGGDSALDWALALRKVAARVTLHRREGFRAADGTVAEMREAVAAGEM
 DFVVGMLGALRTEGQDGPLTEVEVRTRDGSTVLPADDELVALYGLVSEPGPIAQWDM
 DMRAGRILVDTTTTYESSRRGIFAAGDIAYYTNKQKLILSGFHEAALALRRAYHYAFPE
 KALVHVHTSNNAALKEKLTHA

CPMO11

MTASSPDDAVLDVIVGAGLSGIGAARQLQMRCPGKRYAILEAREAMGGTWDLFRY
 PGIRSDSDMYTLGYRFPKWRGAKAIADGPSIRAYIRETAAEEAGITSHIRFGHRVISAA
 WDSAAACWTVEAERSADRSRLRLRTRLLYVCAGYSSYAEGHRPEFAGEAQFRGSM
 VHPQFWDSELDYAGKRVVVIGSGATAVTLVPAMAKSAAHVMTLQRSPTYIVTRPGE
 DAIAHKLRRVLPERLAYAATRWNKVNLLGMMFFQLARRRPDRVRQRLIGMAAAQLAP
 GYDVGTHFTPRYRPWDQRLCLVPDGDLFRAIRDGRASVTDIDRFTENGIVLASGK
 TLAADVVVVATGLKLNMLGDIALTVDGQPRRPAECLAYKGMMLSDVPNLVLAFGYTN
 ASWTLKADLTAEYVCRLLRHMDRHAHRIAVPRAPADVEPTPFLDFTSGYVQRAATVL
 PRQGHRKPWRVHQNYLKDLLAIRHGRIADGVLQFDVPPPSPLAHRHRDGDQDTTDN
 NNTVPAVAASEVQS

CPMO12

MPSFDYDIVIAGASFAGAACALAAARAGLRVVLERKTDPGDKLHTTGILVKEAIEQT
 WLGRAPADCVHRVEQVRLYSPRLRSIDLSAPGYFHTTDTPLMRWLAAELVRHRV
 ELRLGQAFVAVPRNGGWHVTGVGNTRYLVGADGAKSRVAQCAGLGRVHSFLYG
 VEYEFPLALPEPGALHCFITRRFAPGYIGWIAQNPTGVQAGLALRYLPRHGRAPDL
 DGLLAHVREERAGLPQLAQAATRAGLIPCGGPAFPLARDGVILTGDAAGIVSPVTAG
 GIHYAFAHGESVGGQAVAAHLLRGGPAPEAVATAGAPRFRAKRALRWMFDRFQCDW
 PFDLMLYSPPLRWAAEQIYFHKRRGEDPAPPPRAATSA

CPMO13

MYIEPSNAAIPSAAGTNPLAARSELRDVFTNIARQRQIAESDLRKYLADVDPALLASL
 VHITRDPSLLSRYAPELKMPDKGDLMPGVKVVETLKGFAQVATIPGAIRQEMIETVA
 ALDKPASEQTEYLSGYDLAPDLFLKMAAMVTGEPIDAEFGDMLLEQAGFVRNQPVL
 QRGRPTERIELAIGAGMSGIAAAIQAKDRGFRYRIFDSNNKVGGVWAANDYPGVAV
 DTPATYYSLSYELNPSWSNYYPVGSEYLRYLEGIVEKHNISEFIELESEILKIQWIEED
 QEWELMVVKKGREASRVRATAVMSCGLHLNRPNYPDLQGRETFKGVSIHANRWKH
 DVDLRGKRVIIGTGATGVQVIGKISQEVHDLTVFMRQPMWIIPNQAGDGDIPQSKR
 WARQYLPYFLHWDRLKSYWTYSDKVGYPIVRADPEWAKTHVSISPANDRLMQFCIN
 YINSCFGESESELARKLTPDYAPLGKRSVRDPYDFAPGGFYALSQPNVALETSKLAR
 VVPDGLTVDGKLIELDVIIYATGLTLDWLSPIEVAGRNGVRLSDVWRDNNPSSYLG
 MVPKFPNLFINSGPHTGAGHAGGHNFMAEVDNHYAMECLQLLVEEDARSLEVTQEA
 FDEHNRKLDEKMAGAIWAWERRAHTYYTNQKGRPILPTPWRHVDFWRMTREPMKE
 AFILR

CPMO14

MKTETTDVVIIGAGPAGSVAAGLLRKHGIPVLVIEKETFPRFSIGESLLPQSMAYIEEA
 GMLRAVVEAGFQYKNGAAFARGDQQTAFDFREKFSPGWGTTYQVQRRARFDDVLIR
 EAARQGAEVRFSHLVEDVDVSGAQPEVTVRAPDGSRYSVRAKFLLDASGFGRILPR
 LLQLESPSGFPVRSALFTHVEDRIAPGAFDRNKILISVHPQHPDVVYWTIPFSDGRCS
 QGVVAEKAFLDRYTGTETERLRALVAEPEGLAKLLGNAQWDTPARQIAGYSANVKS
 WNGYALLGNAGEFLDPVFSSGVTFIAFKSASLAAQCLVRQLAGDQDVEQDFAQPL
 KAGVDCFRVFVEAWYEGRFQQLIFHPNANSEIRDMISSILAGYAWDRSNPFVAEPRR
 RLAVLEEFGRV

CPMO15

MLQSIPNTKAVQAPVVVIGTGPVGVRLAQLHRRDPSIPLVIYGNPEWEPYNRVRLS
 SFLAGDASWATLTQGLALPQGAIAARLNCAVVHVDRGAQEIEDSAGRTQRYSHLVL
 ATGSRPRIPDVPQIGQANVYVFRDMNDTQRLLARRIRSRRVVLGGGLLGLEAARAM
 QRFNTEVLVIENNDRLMMQQLDQAAADLLRAHVEAAGIQVLTGAAVKRVHGEAVT
 GITLRTGDVIECDTVVIAAGIKPNTDLALQAGLNVGRGIRVSDRMQTSDPHIYAVGEC
 AEHREQVYGIVAPGLEQAAVAVHAICGGDASYAGSLQATRLKVLDMPVFSVGRVGE
 HDGIGSASTAAYEDKKTGIYRKVLVERGHLIGAMAVGEGADLRRLQDAVSGNHRILP
 WQLWRFRRRTGSLWQEGEASVVNWPASAVVCNCTGVTRGQLGEAAAAGCRSIEQL
 AACTRASTVCGSCKPLLAELAGGAVMREPVRHRPLQGAAAAALALALLALVAPAIP
 FAKTVQLQWRWDELWLVS LYKQVSGFSILGLSVLALLSFRKRWSRLGYGDFAIWR
 VLHVVLGLGVLAGLYAHTGGRLGSQLNLLMASFTGLILAGSMSGLMIAQGHRLPVA
 LVRRWRDNDSTWVHILLFWPVPVLLGFHILKTYF

CPMO16

MTPSATPQARPRLVVVGNGMAGMRTVEELLRLAPDLYDITVFGAEPHGNYNRILLSP
 VLAGEKTVADIMLNTREWYAEHGIELLAGDPVVAIDRPRRIVRSASGREVRYDRLLLA
 TGSKPFILPVPQGHQLDGVIAFRDIQDVETMQQAARDHRHAVVIGGGLLGLEAANGLL
 RQGM DVTVVHLPDSL MERQLDKPAATLLQGALERKGLRFLPGAQTAELGIDRVTVG
 RFKDGSEIPADLVVMTAGVVRPNIELAAGAGLHCERAIVDDTLQTYDPRIYAVGECVQ
 HRQATFGLVAPIWDQARVCAAHLAGAGHRHYVQQATATKLKVTGVDLYSAGDFIGA

EGSEDLVLRDPRRGVYKRLVLQDGYLVGAVLYGDVQDGPWYFDMIQRRTPVGALR QRLFLGQAQCEALAA
CPMO17 MKVISSPHELLHQFRGRDSDVPADSQHQIVIVGGGAGGLELATALGNKLGKRGLAAV TLIDKTRSHIWKPKLHEIAAGSMDMNVHEVEFLAQSHWHHFTYRVGEMIGLDRERQE VEVAAFIDDEGHQVTPQRRFRYDTLVIAVGSQSNDFGTPGVSENALKLESPADARRF NQRLVNACIRAHAQPAPLEPHQLQVAIIGAGATDVELAAELHRTRTRALVAYGLDRIDP EKDIRLNLEAADRVLPALPARLSAATEVLLRKLGLVELHTSSRVAEVLPPQGVQLADGRI IAAELVVAAGVKAPEHLKELAGLETNRINQLVVRQTLQTTLDDNIFALGDCAACPWP EKNSIVPPRAQAAHQASHIVGQVLRRLRGKPLRQYRYRDFGSLVSLGEYSTVGNM MGGLSGGNLMIEGSFARLMYLSLYKMHELALHGFPKVALDTLARLITRRTEPHVKLH
HAPMO 4-hydroxy acetophenone monooxygenase [<i>Pseudomonas fluorescens</i>] MSAFNTTLPGLDYDDDTLREHLQGADIPTLLLTVAHLTGDLQILKPNWKPSIAMGVAR SGMDLETEAQVREFCLQRLIDFRDSDGQPAPGRPTSDQLHILGTWLMGPVIEPYLPLIA EEAVTAEEDLRAPRWHKDHVASGRDFKVVIIIGAGESGMIAALRFKQAGVPPFVIYEKG NDVGGTWRENTYPGCRVDINSFWYSFSFARGIWDDCFAPAPQVFAYMQAVAREHG LYEHIRFNTEVSDAHWDESTQRWQLLYRDSEGQTQVDSNVVVFVAVGQLNRPMIPAI PGIETFKGPMFHSAQWDHDVDWVGKRVGVIGTGASATQFIPQLAQTAELKVFART TNWLLPTPDLHEKISDSCKWLLAHVPHYSLWYRVAMAMPQSVGFLEDVMVDVGY PTELAVSARNDRLRQDISAWMEPQFADRPDLREVLIPDSPVGGKRIVRDNGTWISTL KRDNVSMIRQPIEVITPKGICCVDTGTEHEFDLIVYGTGFHASKFLMPINVTGRDGV ALHDVWVGDDARAYLGMTVPQFPNMFCMYGPNTGLVVYSTVIQFSEMTASYIVDAV RLLLEGGHQSMEVKTPVFESYNQRVDEGNALRAWGFSKVNSWYKNSKGRVTQNF PFTAVEFWQRTHSVEPTDYQLG
CtCPMO cyclopentanone monooxygenase [<i>Comamonas testosteroni</i>] MTTMTTMTTEQLGMNNSVNDKLDVLLIGAGFTGLYQLYHLRKLGYKVHLVDAGADIG GIWHWNCYPGARVDTHCQIQYSIPELWQEFNWKELFPNWAQMREYFHFADKKLD LSKDISFNTRVQSAVFDEGTREWTVRSIGHQPIQARFVIANLGFASPSTPNVDGIET FKGQWYHTALWPQEGVNMAGKRVAIIGTGSSGVQVAQEAALDAKQVTVYQRTPNL ALPMHQKQLSAEDNLRMKPELPAAFERRGKCFAGDFDFIAKNATELSAAERTEILE ELWNAGGFRYWLANFQDYLFDDKANDYVYEFWRDKVRARIKDPKVAEKLAPMKKP HPYGAKRPSLEQWYYEIFNQNNVTLVDVNETPVLRITEKGIVTAEGEAFFDLIVFATG FDAVTGGLTSIDFRNNQGSFKDVWSDGIRTQLGVATAGFPNLLFGYGPQSPAGFC NGPSSAEYQGDLLIQLMNYLRDNNISRIEAQSEAQEEWSKLIADFWDSSLFPRAKSW YQGSNIPGKKVESLNFPLGLPTYISKFNESAEGYAGFSLAS
CHMO cyclohexanone monooxygenase [<i>Acinetobacter calcoaceticus</i>] MTAQMKTNASAKKAVNSPNHIFDTFIVGAGISGIATAIRLDQVGYTNYKIIKASRVGG TWRENTYPGCGCDVPSALYSYSFAPSAKWSHLFARQPEILSYLEDVSNEFNVASKIE FNNELLNAAWDDSRHLWVLDTTTGQYLSKTVIFATGPITEAQIPRLEGLDTFKGEMFH SAKWNHDYDLTGKRIAVIGTGASAIQFVPQIQPKAKELLVFQRTAPWVLPKPDTDLGE FSKSVIKYPYQASWRKSVALTNLAINFLRNPLALKPVNVLGKQLLKLQINDPVLRK NVTNPFDIGCKRILFANNYPALQSPNTTLPHGLVKVEGNTVVAANGERHEVDVIV

GTGFEVSHPPIGKR VHNEKGQRLQDLWKDSSPEAYLGTNIENVPNAFLILGPNVLVY
DSFIGLAEAQLDYIVDGLLKIKNKGISKFNVKPEVIKKHNDLVQKHLKTTVFNSGGCKS
YYLDANGRNFAAWPWSLKKLKQRLKSVDLKDYEVTYQTTKTH

PAMO phenylacetone monooxygenase [*Thermobifida fusca*]

MAGQTTVDSRRQPPEEVDVLVVGAGFSGLYALYRLRELGRSVHVIETAGDVGGVW
YWNRYPGARCDIESIEYCYSFSEEVLQEWNWTERYASQPEILRYINFVADKFDLRSGI
TFHTTVTAAAFDEATNTWTVDTNHGDRIRARYLIMASGQLSVPQLPNFPGLKDFAGN
LYHTGNWPHEPVD FSGQRVGVIGTGSSGIQVSPQIAKQAAELFVFQRTPHFAVPAR
NAPLDPEFLADLKKRYAEFREESRNTPPGGTHRYQGPKSALEVSDEELVETLERYWQ
EGGPDILAAAYRDILRDRDANERVAEFIRNKIRNTVRDPEVAERLVPKGYPPFGTKRLILE
IDYYEMFNVDNVHLVDTLSAPIETITPRGVRTSERYELDSLVLATGFDALTGALFKIDI
RGVGNVALKEKWAAGPRTYLGLSTAGFPNLFFIAGPGSPSALS NMLVSIEQHVEWVT
DHIAYMFKNGLTRSEAVLEKEDEWVEHVNEIADETLYPMTASWYTGANVPGKPRVF
MLYVGGFHRYRQICDEVAACKGYEGFVLT

Appendix B The nucleotide encoding regions in *C.necator* H16.

Table 42. The nucleotide encoding regions in chromosome 1 of *C.necator* H16.

WP_197540020.1	WP_051398532.1	WP_041687111.1	WP_011616133.1	WP_011615870.1
WP_197540019.1	WP_051398530.1	WP_041687105.1	WP_011616132.1	WP_011615869.1
WP_197540018.1	WP_051398528.1	WP_041687101.1	WP_011616130.1	WP_011615868.1
WP_197540017.1	WP_051398523.1	WP_041687099.1	WP_011616129.1	WP_011615867.1
WP_197540016.1	WP_051398520.1	WP_041687097.1	WP_011616128.1	WP_011615866.1
WP_197540015.1	WP_051398516.1	WP_041687095.1	WP_011616127.1	WP_011615865.1
WP_197540014.1	WP_051398513.1	WP_041687089.1	WP_011616126.1	WP_011615864.1
WP_197540013.1	WP_051398511.1	WP_041687086.1	WP_011616125.1	WP_011615863.1
WP_197540012.1	WP_051398508.1	WP_041687084.1	WP_011616124.1	WP_011615861.1
WP_197540011.1	WP_051398506.1	WP_041687082.1	WP_011616122.1	WP_011615859.1
WP_197540010.1	WP_051398504.1	WP_037025387.1	WP_011616121.1	WP_011615858.1
WP_197540009.1	WP_051398502.1	WP_037025383.1	WP_011616120.1	WP_011615857.1
WP_193385848.1	WP_051398501.1	WP_037025353.1	WP_011616119.1	WP_011615856.1
WP_193385847.1	WP_051398499.1	WP_037025321.1	WP_011616114.1	WP_011615855.1
WP_193385846.1	WP_051398498.1	WP_037025319.1	WP_011616113.1	WP_011615854.1
WP_176579674.1	WP_051398496.1	WP_037025304.1	WP_011616112.1	WP_011615853.1
WP_174549483.1	WP_051398494.1	WP_037025299.1	WP_011616111.1	WP_011615852.1
WP_174549480.1	WP_051398493.1	WP_037025292.1	WP_011616109.1	WP_011615851.1
WP_174549479.1	WP_051398492.1	WP_037025284.1	WP_011616108.1	WP_011615849.1
WP_174549478.1	WP_051398491.1	WP_037025280.1	WP_011616107.1	WP_011615848.1
WP_174549476.1	WP_051398490.1	WP_037025273.1	WP_011616106.1	WP_011615846.1
WP_174549475.1	WP_051398489.1	WP_037025144.1	WP_011616104.1	WP_011615845.1
WP_174549472.1	WP_051398486.1	WP_037025083.1	WP_011616103.1	WP_011615844.1
WP_174549471.1	WP_051398481.1	WP_037024973.1	WP_011616102.1	WP_011615843.1
WP_174549468.1	WP_051398480.1	WP_037024946.1	WP_011616100.1	WP_011615842.1
WP_174549467.1	WP_051398479.1	WP_037024914.1	WP_011616099.1	WP_011615841.1
WP_174549465.1	WP_051398476.1	WP_037024891.1	WP_011616097.1	WP_011615840.1
WP_174549460.1	WP_051398474.1	WP_037024885.1	WP_011616096.1	WP_011615839.1
WP_174549456.1	WP_051398473.1	WP_037024644.1	WP_011616095.1	WP_011615838.1
WP_174549455.1	WP_051398472.1	WP_037024588.1	WP_011616094.1	WP_011615837.1
WP_174549453.1	WP_051398466.1	WP_037024479.1	WP_011616092.1	WP_011615836.1
WP_174549452.1	WP_051398464.1	WP_037024475.1	WP_011616091.1	WP_011615835.1
WP_174549451.1	WP_051398462.1	WP_037024469.1	WP_011616090.1	WP_011615834.1
WP_174549450.1	WP_051398461.1	WP_037024407.1	WP_011616089.1	WP_011615833.1
WP_172583518.1	WP_051398460.1	WP_037024398.1	WP_011616088.1	WP_011615832.1
WP_167686589.1	WP_051398458.1	WP_037024397.1	WP_011616087.1	WP_011615831.1
WP_167686588.1	WP_051398457.1	WP_037024362.1	WP_011616086.1	WP_011615830.1

Appendix B

WP_167686587.1	WP_051398456.1	WP_037024358.1	WP_011616085.1	WP_011615829.1
WP_167686586.1	WP_051398455.1	WP_037024357.1	WP_011616084.1	WP_011615828.1
WP_167686585.1	WP_051398454.1	WP_037024296.1	WP_011616083.1	WP_011615827.1
WP_167686584.1	WP_051398453.1	WP_037024292.1	WP_011616082.1	WP_011615826.1
WP_167686583.1	WP_051398452.1	WP_037024113.1	WP_011616081.1	WP_011615825.1
WP_167686582.1	WP_041687850.1	WP_037024101.1	WP_011616079.1	WP_011615824.1
WP_167686581.1	WP_041687849.1	WP_037024094.1	WP_011616078.1	WP_011615823.1
WP_167686580.1	WP_041687848.1	WP_037024085.1	WP_011616077.1	WP_011615822.1
WP_167686579.1	WP_041687847.1	WP_037024065.1	WP_011616076.1	WP_011615820.1
WP_167686578.1	WP_041687846.1	WP_037023964.1	WP_011616075.1	WP_011615817.1
WP_167686577.1	WP_041687845.1	WP_037023942.1	WP_011616073.1	WP_011615816.1
WP_167686576.1	WP_041687844.1	WP_037023915.1	WP_011616072.1	WP_011615815.1
WP_167686575.1	WP_041687843.1	WP_037023863.1	WP_011616071.1	WP_011615814.1
WP_167686574.1	WP_041687841.1	WP_037023859.1	WP_011616070.1	WP_011615813.1
WP_162166398.1	WP_041687839.1	WP_037023846.1	WP_011616066.1	WP_011615812.1
WP_162166397.1	WP_041687838.1	WP_037023841.1	WP_011616064.1	WP_011615811.1
WP_162166395.1	WP_041687832.1	WP_037023832.1	WP_011616061.1	WP_011615810.1
WP_162166394.1	WP_041687829.1	WP_037023640.1	WP_011616060.1	WP_011615809.1
WP_157886292.1	WP_041687823.1	WP_037023635.1	WP_011616059.1	WP_011615807.1
WP_157886291.1	WP_041687815.1	WP_037023622.1	WP_011616058.1	WP_011615806.1
WP_157886290.1	WP_041687813.1	WP_037023614.1	WP_011616057.1	WP_011615805.1
WP_157886289.1	WP_041687812.1	WP_037023579.1	WP_011616054.1	WP_011615804.1
WP_157886288.1	WP_041687810.1	WP_037023529.1	WP_011616052.1	WP_011615802.1
WP_157886286.1	WP_041687807.1	WP_037023519.1	WP_011616051.1	WP_011615801.1
WP_157886285.1	WP_041687806.1	WP_037023515.1	WP_011616050.1	WP_011615799.1
WP_157886283.1	WP_041687802.1	WP_037023390.1	WP_011616049.1	WP_011615797.1
WP_157886281.1	WP_041687792.1	WP_037023387.1	WP_011616048.1	WP_011615796.1
WP_157886280.1	WP_041687790.1	WP_037023310.1	WP_011616046.1	WP_011615795.1
WP_157886278.1	WP_041687771.1	WP_037023308.1	WP_011616043.1	WP_011615794.1
WP_157886276.1	WP_041687751.1	WP_035821236.1	WP_011616041.1	WP_011615792.1
WP_155737094.1	WP_041687741.1	WP_035818789.1	WP_011616040.1	WP_011615789.1
WP_155737093.1	WP_041687722.1	WP_035817912.1	WP_011616038.1	WP_011615788.1
WP_155737088.1	WP_041687720.1	WP_035815700.1	WP_011616037.1	WP_011615787.1
WP_155737086.1	WP_041687716.1	WP_026201385.1	WP_011616036.1	WP_011615786.1
WP_155737085.1	WP_041687714.1	WP_026201125.1	WP_011616035.1	WP_011615785.1
WP_155737083.1	WP_041687712.1	WP_026200447.1	WP_011616034.1	WP_011615784.1
WP_155737082.1	WP_041687694.1	WP_022539889.1	WP_011616033.1	WP_011615783.1
WP_155737081.1	WP_041687681.1	WP_013956405.1	WP_011616032.1	WP_011615782.1
WP_155737079.1	WP_041687679.1	WP_011616294.1	WP_011616031.1	WP_011615781.1
WP_155737068.1	WP_041687677.1	WP_011616293.1	WP_011616027.1	WP_011615780.1
WP_153018154.1	WP_041687675.1	WP_011616292.1	WP_011616026.1	WP_011615779.1
WP_136227835.1	WP_041687668.1	WP_011616291.1	WP_011616025.1	WP_011615778.1

Appendix B

WP_136227834.1	WP_041687666.1	WP_011616290.1	WP_011616023.1	WP_011615777.1
WP_136227833.1	WP_041687656.1	WP_011616289.1	WP_011616022.1	WP_011615776.1
WP_136227831.1	WP_041687641.1	WP_011616288.1	WP_011616021.1	WP_011615774.1
WP_136227830.1	WP_041687635.1	WP_011616287.1	WP_011616020.1	WP_011615773.1
WP_136227829.1	WP_041687633.1	WP_011616286.1	WP_011616018.1	WP_011615772.1
WP_136227828.1	WP_041687625.1	WP_011616285.1	WP_011616017.1	WP_011615771.1
WP_136227827.1	WP_041687623.1	WP_011616284.1	WP_011616015.1	WP_011615769.1
WP_136227826.1	WP_041687615.1	WP_011616283.1	WP_011616014.1	WP_011615767.1
WP_136227825.1	WP_041687606.1	WP_011616282.1	WP_011616013.1	WP_011615765.1
WP_136227824.1	WP_041687602.1	WP_011616281.1	WP_011616012.1	WP_011615763.1
WP_136227823.1	WP_041687594.1	WP_011616279.1	WP_011616011.1	WP_011615762.1
WP_136227822.1	WP_041687590.1	WP_011616278.1	WP_011616010.1	WP_011615761.1
WP_136227821.1	WP_041687579.1	WP_011616276.1	WP_011616009.1	WP_011615760.1
WP_136227820.1	WP_041687575.1	WP_011616275.1	WP_011616007.1	WP_011615759.1
WP_136227819.1	WP_041687572.1	WP_011616273.1	WP_011616006.1	WP_011615758.1
WP_136227818.1	WP_041687570.1	WP_011616272.1	WP_011616005.1	WP_011615757.1
WP_136227817.1	WP_041687562.1	WP_011616271.1	WP_011616004.1	WP_011615756.1
WP_136227816.1	WP_041687559.1	WP_011616270.1	WP_011616003.1	WP_011615753.1
WP_136227814.1	WP_041687557.1	WP_011616269.1	WP_011616002.1	WP_011615752.1
WP_136227813.1	WP_041687550.1	WP_011616268.1	WP_011616001.1	WP_011615751.1
WP_136227812.1	WP_041687546.1	WP_011616267.1	WP_011616000.1	WP_011615750.1
WP_136227811.1	WP_041687544.1	WP_011616266.1	WP_011615999.1	WP_011615749.1
WP_136227810.1	WP_041687541.1	WP_011616265.1	WP_011615998.1	WP_011615748.1
WP_136227809.1	WP_041687532.1	WP_011616264.1	WP_011615997.1	WP_011615747.1
WP_136227808.1	WP_041687523.1	WP_011616263.1	WP_011615996.1	WP_011615746.1
WP_136227807.1	WP_041687516.1	WP_011616262.1	WP_011615995.1	WP_011615745.1
WP_136227806.1	WP_041687510.1	WP_011616261.1	WP_011615994.1	WP_011615744.1
WP_136227805.1	WP_041687507.1	WP_011616260.1	WP_011615993.1	WP_011615743.1
WP_136227804.1	WP_041687497.1	WP_011616259.1	WP_011615991.1	WP_011615742.1
WP_136227803.1	WP_041687494.1	WP_011616258.1	WP_011615990.1	WP_011615741.1
WP_136227801.1	WP_041687490.1	WP_011616257.1	WP_011615989.1	WP_011615737.1
WP_136227800.1	WP_041687488.1	WP_011616255.1	WP_011615988.1	WP_011615735.1
WP_136227798.1	WP_041687478.1	WP_011616254.1	WP_011615987.1	WP_011615734.1
WP_136227797.1	WP_041687474.1	WP_011616252.1	WP_011615986.1	WP_011615733.1
WP_136227796.1	WP_041687472.1	WP_011616251.1	WP_011615985.1	WP_011615732.1
WP_136227794.1	WP_041687468.1	WP_011616249.1	WP_011615984.1	WP_011615731.1
WP_136227793.1	WP_041687464.1	WP_011616248.1	WP_011615983.1	WP_011615730.1
WP_136227792.1	WP_041687462.1	WP_011616246.1	WP_011615981.1	WP_011615728.1
WP_136227791.1	WP_041687459.1	WP_011616245.1	WP_011615979.1	WP_011615727.1
WP_136227784.1	WP_041687455.1	WP_011616244.1	WP_011615977.1	WP_011615726.1
WP_136227783.1	WP_041687450.1	WP_011616243.1	WP_011615976.1	WP_011615725.1
WP_136227780.1	WP_041687443.1	WP_011616240.1	WP_011615975.1	WP_011615724.1

Appendix B

WP_136227779.1	WP_041687440.1	WP_011616239.1	WP_011615971.1	WP_011615723.1
WP_136227778.1	WP_041687438.1	WP_011616238.1	WP_011615969.1	WP_011615721.1
WP_136227777.1	WP_041687434.1	WP_011616237.1	WP_011615968.1	WP_011615720.1
WP_136227776.1	WP_041687431.1	WP_011616236.1	WP_011615967.1	WP_011615719.1
WP_136227774.1	WP_041687425.1	WP_011616233.1	WP_011615966.1	WP_011615718.1
WP_136227772.1	WP_041687420.1	WP_011616232.1	WP_011615965.1	WP_011615717.1
WP_136227769.1	WP_041687416.1	WP_011616229.1	WP_011615962.1	WP_011615716.1
WP_136227767.1	WP_041687412.1	WP_011616226.1	WP_011615961.1	WP_011615715.1
WP_136227766.1	WP_041687406.1	WP_011616225.1	WP_011615959.1	WP_011615714.1
WP_136227764.1	WP_041687401.1	WP_011616223.1	WP_011615958.1	WP_011615713.1
WP_136227763.1	WP_041687398.1	WP_011616222.1	WP_011615957.1	WP_011615711.1
WP_136227762.1	WP_041687384.1	WP_011616221.1	WP_011615956.1	WP_011615710.1
WP_136227761.1	WP_041687381.1	WP_011616220.1	WP_011615955.1	WP_011615709.1
WP_136227759.1	WP_041687360.1	WP_011616216.1	WP_011615953.1	WP_011615708.1
WP_136227758.1	WP_041687358.1	WP_011616213.1	WP_011615951.1	WP_011615706.1
WP_136227756.1	WP_041687356.1	WP_011616212.1	WP_011615950.1	WP_011615702.1
WP_136227752.1	WP_041687352.1	WP_011616211.1	WP_011615949.1	WP_011615701.1
WP_136227748.1	WP_041687341.1	WP_011616210.1	WP_011615948.1	WP_011615700.1
WP_136227746.1	WP_041687332.1	WP_011616209.1	WP_011615947.1	WP_011615699.1
WP_136227744.1	WP_041687326.1	WP_011616208.1	WP_011615946.1	WP_011615698.1
WP_136227743.1	WP_041687318.1	WP_011616207.1	WP_011615944.1	WP_011615697.1
WP_136227742.1	WP_041687316.1	WP_011616206.1	WP_011615943.1	WP_011615696.1
WP_136227740.1	WP_041687310.1	WP_011616205.1	WP_011615942.1	WP_011615695.1
WP_136227739.1	WP_041687301.1	WP_011616204.1	WP_011615941.1	WP_011615692.1
WP_136227735.1	WP_041687296.1	WP_011616202.1	WP_011615939.1	WP_011615691.1
WP_136227734.1	WP_041687294.1	WP_011616201.1	WP_011615937.1	WP_011615686.1
WP_136227733.1	WP_041687290.1	WP_011616200.1	WP_011615936.1	WP_011615684.1
WP_136227730.1	WP_041687285.1	WP_011616199.1	WP_011615935.1	WP_011615676.1
WP_136227729.1	WP_041687282.1	WP_011616198.1	WP_011615933.1	WP_011615674.1
WP_136227728.1	WP_041687280.1	WP_011616197.1	WP_011615932.1	WP_011615673.1
WP_099045624.1	WP_041687278.1	WP_011616196.1	WP_011615930.1	WP_011615672.1
WP_099045608.1	WP_041687274.1	WP_011616195.1	WP_011615929.1	WP_011615671.1
WP_099045607.1	WP_041687270.1	WP_011616194.1	WP_011615928.1	WP_011615670.1
WP_099045602.1	WP_041687266.1	WP_011616193.1	WP_011615927.1	WP_011615669.1
WP_099045597.1	WP_041687263.1	WP_011616191.1	WP_011615926.1	WP_011615668.1
WP_099045594.1	WP_041687261.1	WP_011616190.1	WP_011615925.1	WP_011615667.1
WP_082236111.1	WP_041687259.1	WP_011616188.1	WP_011615924.1	WP_011615666.1
WP_081623904.1	WP_041687253.1	WP_011616187.1	WP_011615923.1	WP_011615663.1
WP_081225812.1	WP_041687249.1	WP_011616186.1	WP_011615922.1	WP_011615662.1
WP_081225811.1	WP_041687246.1	WP_011616185.1	WP_011615920.1	WP_011615661.1
WP_081225809.1	WP_041687243.1	WP_011616184.1	WP_011615919.1	WP_011615660.1
WP_081225807.1	WP_041687238.1	WP_011616183.1	WP_011615918.1	WP_011615659.1

Appendix B

WP_081225805.1	WP_041687231.1	WP_011616182.1	WP_011615917.1	WP_011615657.1
WP_081225801.1	WP_041687227.1	WP_011616181.1	WP_011615916.1	WP_011615656.1
WP_081225799.1	WP_041687224.1	WP_011616180.1	WP_011615914.1	WP_011615654.1
WP_081225797.1	WP_041687221.1	WP_011616179.1	WP_011615912.1	WP_011615652.1
WP_081225790.1	WP_041687219.1	WP_011616178.1	WP_011615911.1	WP_011615651.1
WP_081225789.1	WP_041687215.1	WP_011616177.1	WP_011615908.1	WP_011615650.1
WP_081225788.1	WP_041687213.1	WP_011616176.1	WP_011615907.1	WP_011615649.1
WP_081225787.1	WP_041687211.1	WP_011616175.1	WP_011615906.1	WP_011615648.1
WP_081225777.1	WP_041687206.1	WP_011616173.1	WP_011615905.1	WP_011615647.1
WP_081225771.1	WP_041687204.1	WP_011616171.1	WP_011615904.1	WP_011615645.1
WP_081225769.1	WP_041687200.1	WP_011616169.1	WP_011615903.1	WP_011615644.1
WP_081225768.1	WP_041687198.1	WP_011616168.1	WP_011615902.1	WP_011615643.1
WP_081225763.1	WP_041687196.1	WP_011616167.1	WP_011615901.1	WP_011615642.1
WP_081225759.1	WP_041687194.1	WP_011616166.1	WP_011615900.1	WP_011615641.1
WP_081225757.1	WP_041687192.1	WP_011616165.1	WP_011615898.1	WP_011615640.1
WP_081225755.1	WP_041687185.1	WP_011616164.1	WP_011615897.1	WP_011615639.1
WP_081225754.1	WP_041687181.1	WP_011616163.1	WP_011615895.1	WP_011615638.1
WP_081225753.1	WP_041687179.1	WP_011616162.1	WP_011615893.1	WP_011615636.1
WP_081050391.1	WP_041687174.1	WP_011616161.1	WP_011615892.1	WP_011615635.1
WP_081050386.1	WP_041687171.1	WP_011616159.1	WP_011615890.1	WP_011615634.1
WP_081050382.1	WP_041687169.1	WP_011616158.1	WP_011615889.1	WP_011615632.1
WP_081050336.1	WP_041687167.1	WP_011616157.1	WP_011615888.1	WP_011615631.1
WP_081050286.1	WP_041687165.1	WP_011616156.1	WP_011615886.1	WP_011615630.1
WP_081050276.1	WP_041687163.1	WP_011616153.1	WP_011615885.1	WP_011615629.1
WP_081050274.1	WP_041687160.1	WP_011616151.1	WP_011615884.1	WP_011615628.1
WP_081050270.1	WP_041687157.1	WP_011616149.1	WP_011615883.1	WP_011615627.1
WP_081050262.1	WP_041687155.1	WP_011616148.1	WP_011615882.1	WP_011615626.1
WP_081050209.1	WP_041687152.1	WP_011616147.1	WP_011615881.1	WP_011615625.1
WP_081050194.1	WP_041687146.1	WP_011616146.1	WP_011615880.1	WP_011615623.1
WP_081050143.1	WP_041687143.1	WP_011616145.1	WP_011615879.1	WP_011615619.1
WP_078195944.1	WP_041687141.1	WP_011616142.1	WP_011615878.1	WP_011615618.1
WP_063834060.1	WP_041687139.1	WP_011616141.1	WP_011615877.1	WP_011615617.1
WP_063834059.1	WP_041687137.1	WP_011616140.1	WP_011615876.1	WP_011615616.1
WP_062804606.1	WP_041687131.1	WP_011616139.1	WP_011615875.1	WP_011615615.1
WP_051398540.1	WP_041687124.1	WP_011616138.1	WP_011615874.1	WP_011615614.1
WP_051398537.1	WP_041687119.1	WP_011616137.1	WP_011615873.1	WP_011615613.1
WP_051398533.1	WP_041687115.1	WP_011616135.1	WP_011615871.1	WP_011615612.1

WP_011615611.1	WP_011615332.1	WP_011615075.1	WP_011614817.1	WP_011614543.1
WP_011615610.1	WP_011615331.1	WP_011615074.1	WP_011614816.1	WP_011614542.1
WP_011615609.1	WP_011615330.1	WP_011615073.1	WP_011614815.1	WP_011614541.1

Appendix B

WP_011615608.1	WP_011615329.1	WP_011615072.1	WP_011614811.1	WP_011614539.1
WP_011615606.1	WP_011615328.1	WP_011615071.1	WP_011614810.1	WP_011614538.1
WP_011615605.1	WP_011615327.1	WP_011615070.1	WP_011614808.1	WP_011614537.1
WP_011615604.1	WP_011615326.1	WP_011615069.1	WP_011614807.1	WP_011614536.1
WP_011615602.1	WP_011615324.1	WP_011615068.1	WP_011614806.1	WP_011614535.1
WP_011615601.1	WP_011615322.1	WP_011615067.1	WP_011614805.1	WP_011614534.1
WP_011615600.1	WP_011615321.1	WP_011615066.1	WP_011614804.1	WP_011614533.1
WP_011615599.1	WP_011615320.1	WP_011615064.1	WP_011614803.1	WP_011614532.1
WP_011615598.1	WP_011615319.1	WP_011615063.1	WP_011614802.1	WP_011614531.1
WP_011615597.1	WP_011615318.1	WP_011615062.1	WP_011614801.1	WP_011614530.1
WP_011615596.1	WP_011615317.1	WP_011615061.1	WP_011614800.1	WP_011614529.1
WP_011615595.1	WP_011615314.1	WP_011615060.1	WP_011614798.1	WP_011614527.1
WP_011615594.1	WP_011615312.1	WP_011615058.1	WP_011614797.1	WP_011614526.1
WP_011615591.1	WP_011615311.1	WP_011615057.1	WP_011614796.1	WP_011614524.1
WP_011615590.1	WP_011615310.1	WP_011615056.1	WP_011614795.1	WP_011614522.1
WP_011615588.1	WP_011615308.1	WP_011615055.1	WP_011614794.1	WP_011614521.1
WP_011615586.1	WP_011615307.1	WP_011615054.1	WP_011614793.1	WP_011614520.1
WP_011615585.1	WP_011615306.1	WP_011615053.1	WP_011614789.1	WP_011614519.1
WP_011615584.1	WP_011615304.1	WP_011615052.1	WP_011614788.1	WP_011614518.1
WP_011615583.1	WP_011615303.1	WP_011615051.1	WP_011614787.1	WP_011614517.1
WP_011615582.1	WP_011615302.1	WP_011615050.1	WP_011614786.1	WP_011614516.1
WP_011615581.1	WP_011615301.1	WP_011615049.1	WP_011614785.1	WP_011614515.1
WP_011615580.1	WP_011615300.1	WP_011615048.1	WP_011614784.1	WP_011614514.1
WP_011615579.1	WP_011615299.1	WP_011615047.1	WP_011614782.1	WP_011614512.1
WP_011615577.1	WP_011615298.1	WP_011615046.1	WP_011614780.1	WP_011614510.1
WP_011615576.1	WP_011615296.1	WP_011615045.1	WP_011614779.1	WP_011614509.1
WP_011615575.1	WP_011615295.1	WP_011615044.1	WP_011614778.1	WP_011614508.1
WP_011615574.1	WP_011615294.1	WP_011615043.1	WP_011614775.1	WP_011614507.1
WP_011615573.1	WP_011615293.1	WP_011615041.1	WP_011614774.1	WP_011614506.1
WP_011615572.1	WP_011615291.1	WP_011615040.1	WP_011614766.1	WP_011614505.1
WP_011615571.1	WP_011615290.1	WP_011615037.1	WP_011614765.1	WP_011614504.1
WP_011615570.1	WP_011615289.1	WP_011615036.1	WP_011614764.1	WP_011614503.1
WP_011615569.1	WP_011615286.1	WP_011615034.1	WP_011614763.1	WP_011614502.1
WP_011615568.1	WP_011615285.1	WP_011615033.1	WP_011614762.1	WP_011614501.1
WP_011615567.1	WP_011615284.1	WP_011615032.1	WP_011614761.1	WP_011614498.1
WP_011615565.1	WP_011615282.1	WP_011615031.1	WP_011614760.1	WP_011614497.1
WP_011615564.1	WP_011615281.1	WP_011615030.1	WP_011614759.1	WP_011614496.1
WP_011615563.1	WP_011615280.1	WP_011615029.1	WP_011614758.1	WP_011614495.1
WP_011615561.1	WP_011615279.1	WP_011615028.1	WP_011614757.1	WP_011614494.1
WP_011615558.1	WP_011615276.1	WP_011615027.1	WP_011614756.1	WP_011614493.1
WP_011615557.1	WP_011615274.1	WP_011615025.1	WP_011614751.1	WP_011614492.1
WP_011615555.1	WP_011615273.1	WP_011615024.1	WP_011614750.1	WP_011614491.1

Appendix B

WP_011615554.1	WP_011615272.1	WP_011615022.1	WP_011614749.1	WP_011614490.1
WP_011615553.1	WP_011615271.1	WP_011615021.1	WP_011614748.1	WP_011614488.1
WP_011615551.1	WP_011615270.1	WP_011615018.1	WP_011614747.1	WP_011614487.1
WP_011615550.1	WP_011615268.1	WP_011615017.1	WP_011614746.1	WP_011614486.1
WP_011615547.1	WP_011615267.1	WP_011615016.1	WP_011614745.1	WP_011614484.1
WP_011615546.1	WP_011615265.1	WP_011615015.1	WP_011614743.1	WP_011614483.1
WP_011615545.1	WP_011615264.1	WP_011615010.1	WP_011614742.1	WP_011614482.1
WP_011615544.1	WP_011615262.1	WP_011615009.1	WP_011614741.1	WP_011614481.1
WP_011615543.1	WP_011615261.1	WP_011615007.1	WP_011614739.1	WP_011614480.1
WP_011615541.1	WP_011615260.1	WP_011615006.1	WP_011614738.1	WP_011614479.1
WP_011615540.1	WP_011615259.1	WP_011615004.1	WP_011614737.1	WP_011614478.1
WP_011615539.1	WP_011615258.1	WP_011615003.1	WP_011614735.1	WP_011614477.1
WP_011615538.1	WP_011615257.1	WP_011615002.1	WP_011614734.1	WP_011614476.1
WP_011615537.1	WP_011615256.1	WP_011615001.1	WP_011614732.1	WP_011614475.1
WP_011615535.1	WP_011615255.1	WP_011615000.1	WP_011614731.1	WP_011614472.1
WP_011615533.1	WP_011615254.1	WP_011614998.1	WP_011614730.1	WP_011614470.1
WP_011615532.1	WP_011615253.1	WP_011614997.1	WP_011614729.1	WP_011614467.1
WP_011615531.1	WP_011615251.1	WP_011614996.1	WP_011614728.1	WP_011614466.1
WP_011615529.1	WP_011615250.1	WP_011614995.1	WP_011614727.1	WP_011614463.1
WP_011615528.1	WP_011615249.1	WP_011614994.1	WP_011614725.1	WP_011614462.1
WP_011615525.1	WP_011615247.1	WP_011614993.1	WP_011614724.1	WP_011614461.1
WP_011615523.1	WP_011615246.1	WP_011614992.1	WP_011614723.1	WP_011614460.1
WP_011615522.1	WP_011615243.1	WP_011614991.1	WP_011614721.1	WP_011614459.1
WP_011615521.1	WP_011615242.1	WP_011614990.1	WP_011614718.1	WP_011614458.1
WP_011615520.1	WP_011615241.1	WP_011614989.1	WP_011614716.1	WP_011614457.1
WP_011615519.1	WP_011615240.1	WP_011614988.1	WP_011614715.1	WP_011614456.1
WP_011615518.1	WP_011615239.1	WP_011614987.1	WP_011614708.1	WP_011614454.1
WP_011615517.1	WP_011615238.1	WP_011614986.1	WP_011614707.1	WP_011614453.1
WP_011615516.1	WP_011615237.1	WP_011614985.1	WP_011614706.1	WP_011614452.1
WP_011615514.1	WP_011615236.1	WP_011614983.1	WP_011614705.1	WP_011614451.1
WP_011615513.1	WP_011615235.1	WP_011614982.1	WP_011614704.1	WP_011614448.1
WP_011615511.1	WP_011615234.1	WP_011614981.1	WP_011614703.1	WP_011614447.1
WP_011615510.1	WP_011615233.1	WP_011614979.1	WP_011614702.1	WP_011614446.1
WP_011615509.1	WP_011615232.1	WP_011614978.1	WP_011614701.1	WP_011614445.1
WP_011615506.1	WP_011615231.1	WP_011614977.1	WP_011614700.1	WP_011614443.1
WP_011615505.1	WP_011615229.1	WP_011614975.1	WP_011614699.1	WP_011614442.1
WP_011615503.1	WP_011615228.1	WP_011614974.1	WP_011614697.1	WP_011614441.1
WP_011615502.1	WP_011615227.1	WP_011614973.1	WP_011614696.1	WP_011614440.1
WP_011615501.1	WP_011615226.1	WP_011614972.1	WP_011614695.1	WP_011614439.1
WP_011615498.1	WP_011615225.1	WP_011614971.1	WP_011614694.1	WP_011614438.1
WP_011615497.1	WP_011615222.1	WP_011614970.1	WP_011614692.1	WP_011614436.1
WP_011615496.1	WP_011615221.1	WP_011614969.1	WP_011614691.1	WP_011614435.1

Appendix B

WP_011615493.1	WP_011615220.1	WP_011614967.1	WP_011614690.1	WP_011614434.1
WP_011615492.1	WP_011615219.1	WP_011614966.1	WP_011614689.1	WP_011614433.1
WP_011615491.1	WP_011615218.1	WP_011614965.1	WP_011614688.1	WP_011614432.1
WP_011615490.1	WP_011615217.1	WP_011614964.1	WP_011614685.1	WP_011614431.1
WP_011615489.1	WP_011615216.1	WP_011614962.1	WP_011614684.1	WP_011614430.1
WP_011615488.1	WP_011615212.1	WP_011614961.1	WP_011614683.1	WP_011614429.1
WP_011615487.1	WP_011615210.1	WP_011614960.1	WP_011614682.1	WP_011614428.1
WP_011615485.1	WP_011615209.1	WP_011614959.1	WP_011614680.1	WP_011614427.1
WP_011615484.1	WP_011615208.1	WP_011614958.1	WP_011614679.1	WP_011614426.1
WP_011615482.1	WP_011615206.1	WP_011614957.1	WP_011614678.1	WP_011614425.1
WP_011615481.1	WP_011615205.1	WP_011614954.1	WP_011614676.1	WP_011614424.1
WP_011615480.1	WP_011615204.1	WP_011614952.1	WP_011614675.1	WP_011614423.1
WP_011615478.1	WP_011615203.1	WP_011614950.1	WP_011614674.1	WP_011614421.1
WP_011615477.1	WP_011615202.1	WP_011614949.1	WP_011614673.1	WP_011614420.1
WP_011615476.1	WP_011615200.1	WP_011614948.1	WP_011614672.1	WP_011614419.1
WP_011615475.1	WP_011615198.1	WP_011614947.1	WP_011614671.1	WP_011614418.1
WP_011615473.1	WP_011615197.1	WP_011614946.1	WP_011614670.1	WP_011614417.1
WP_011615472.1	WP_011615196.1	WP_011614945.1	WP_011614669.1	WP_011614415.1
WP_011615471.1	WP_011615195.1	WP_011614944.1	WP_011614668.1	WP_011614414.1
WP_011615469.1	WP_011615194.1	WP_011614943.1	WP_011614666.1	WP_011614413.1
WP_011615468.1	WP_011615193.1	WP_011614942.1	WP_011614665.1	WP_011614412.1
WP_011615467.1	WP_011615192.1	WP_011614941.1	WP_011614663.1	WP_011614411.1
WP_011615466.1	WP_011615191.1	WP_011614939.1	WP_011614662.1	WP_011614408.1
WP_011615464.1	WP_011615190.1	WP_011614938.1	WP_011614660.1	WP_011614406.1
WP_011615463.1	WP_011615189.1	WP_011614936.1	WP_011614658.1	WP_011614405.1
WP_011615462.1	WP_011615188.1	WP_011614935.1	WP_011614657.1	WP_011614404.1
WP_011615461.1	WP_011615187.1	WP_011614931.1	WP_011614656.1	WP_011614403.1
WP_011615459.1	WP_011615185.1	WP_011614930.1	WP_011614655.1	WP_011614402.1
WP_011615458.1	WP_011615184.1	WP_011614929.1	WP_011614654.1	WP_011614401.1
WP_011615456.1	WP_011615183.1	WP_011614927.1	WP_011614653.1	WP_011614400.1
WP_011615455.1	WP_011615182.1	WP_011614926.1	WP_011614651.1	WP_011614399.1
WP_011615454.1	WP_011615181.1	WP_011614924.1	WP_011614650.1	WP_011614398.1
WP_011615451.1	WP_011615180.1	WP_011614922.1	WP_011614649.1	WP_011614397.1
WP_011615450.1	WP_011615179.1	WP_011614921.1	WP_011614648.1	WP_011614396.1
WP_011615449.1	WP_011615178.1	WP_011614920.1	WP_011614647.1	WP_011614395.1
WP_011615447.1	WP_011615176.1	WP_011614918.1	WP_011614646.1	WP_011614391.1
WP_011615444.1	WP_011615174.1	WP_011614917.1	WP_011614644.1	WP_011614389.1
WP_011615442.1	WP_011615173.1	WP_011614916.1	WP_011614642.1	WP_011614388.1
WP_011615441.1	WP_011615169.1	WP_011614915.1	WP_011614641.1	WP_011614387.1
WP_011615438.1	WP_011615168.1	WP_011614914.1	WP_011614639.1	WP_011614386.1
WP_011615436.1	WP_011615167.1	WP_011614912.1	WP_011614638.1	WP_011614384.1
WP_011615435.1	WP_011615166.1	WP_011614911.1	WP_011614635.1	WP_011614383.1

Appendix B

WP_011615434.1	WP_011615165.1	WP_011614910.1	WP_011614634.1	WP_011614382.1
WP_011615433.1	WP_011615164.1	WP_011614909.1	WP_011614633.1	WP_011614381.1
WP_011615432.1	WP_011615163.1	WP_011614907.1	WP_011614632.1	WP_011614380.1
WP_011615431.1	WP_011615162.1	WP_011614906.1	WP_011614631.1	WP_011614379.1
WP_011615429.1	WP_011615161.1	WP_011614905.1	WP_011614630.1	WP_011614378.1
WP_011615428.1	WP_011615160.1	WP_011614904.1	WP_011614628.1	WP_011614377.1
WP_011615425.1	WP_011615159.1	WP_011614903.1	WP_011614627.1	WP_011614376.1
WP_011615423.1	WP_011615156.1	WP_011614902.1	WP_011614625.1	WP_011614375.1
WP_011615422.1	WP_011615155.1	WP_011614900.1	WP_011614624.1	WP_011614371.1
WP_011615421.1	WP_011615154.1	WP_011614899.1	WP_011614623.1	WP_011614370.1
WP_011615420.1	WP_011615152.1	WP_011614898.1	WP_011614622.1	WP_011614369.1
WP_011615419.1	WP_011615151.1	WP_011614897.1	WP_011614621.1	WP_011614368.1
WP_011615417.1	WP_011615150.1	WP_011614896.1	WP_011614618.1	WP_011614366.1
WP_011615416.1	WP_011615149.1	WP_011614895.1	WP_011614616.1	WP_011614365.1
WP_011615415.1	WP_011615148.1	WP_011614894.1	WP_011614614.1	WP_011614364.1
WP_011615414.1	WP_011615147.1	WP_011614892.1	WP_011614613.1	WP_011614363.1
WP_011615413.1	WP_011615145.1	WP_011614891.1	WP_011614611.1	WP_011614362.1
WP_011615412.1	WP_011615144.1	WP_011614889.1	WP_011614610.1	WP_011614361.1
WP_011615410.1	WP_011615143.1	WP_011614888.1	WP_011614609.1	WP_011614359.1
WP_011615409.1	WP_011615142.1	WP_011614887.1	WP_011614608.1	WP_011614358.1
WP_011615408.1	WP_011615141.1	WP_011614886.1	WP_011614607.1	WP_011614355.1
WP_011615407.1	WP_011615140.1	WP_011614885.1	WP_011614603.1	WP_011614354.1
WP_011615406.1	WP_011615139.1	WP_011614884.1	WP_011614602.1	WP_011614351.1
WP_011615405.1	WP_011615138.1	WP_011614883.1	WP_011614601.1	WP_011614350.1
WP_011615402.1	WP_011615136.1	WP_011614882.1	WP_011614599.1	WP_011614347.1
WP_011615401.1	WP_011615135.1	WP_011614879.1	WP_011614598.1	WP_011614346.1
WP_011615400.1	WP_011615134.1	WP_011614878.1	WP_011614597.1	WP_011614344.1
WP_011615398.1	WP_011615132.1	WP_011614877.1	WP_011614596.1	WP_011614343.1
WP_011615397.1	WP_011615131.1	WP_011614876.1	WP_011614595.1	WP_011614342.1
WP_011615396.1	WP_011615130.1	WP_011614875.1	WP_011614593.1	WP_011614341.1
WP_011615395.1	WP_011615128.1	WP_011614872.1	WP_011614592.1	WP_011614340.1
WP_011615392.1	WP_011615127.1	WP_011614871.1	WP_011614591.1	WP_011614338.1
WP_011615390.1	WP_011615125.1	WP_011614870.1	WP_011614590.1	WP_011614336.1
WP_011615389.1	WP_011615124.1	WP_011614869.1	WP_011614588.1	WP_011614335.1
WP_011615388.1	WP_011615123.1	WP_011614868.1	WP_011614587.1	WP_011614334.1
WP_011615387.1	WP_011615122.1	WP_011614867.1	WP_011614586.1	WP_011614332.1
WP_011615384.1	WP_011615121.1	WP_011614866.1	WP_011614584.1	WP_011614331.1
WP_011615383.1	WP_011615120.1	WP_011614865.1	WP_011614582.1	WP_011614330.1
WP_011615382.1	WP_011615119.1	WP_011614864.1	WP_011614581.1	WP_011614329.1
WP_011615381.1	WP_011615118.1	WP_011614863.1	WP_011614580.1	WP_011614328.1
WP_011615380.1	WP_011615117.1	WP_011614861.1	WP_011614579.1	WP_011614326.1
WP_011615377.1	WP_011615116.1	WP_011614860.1	WP_011614578.1	WP_011614325.1

Appendix B

WP_011615376.1	WP_011615115.1	WP_011614859.1	WP_011614576.1	WP_011614324.1
WP_011615374.1	WP_011615112.1	WP_011614858.1	WP_011614574.1	WP_011614322.1
WP_011615373.1	WP_011615110.1	WP_011614857.1	WP_011614573.1	WP_011614321.1
WP_011615372.1	WP_011615108.1	WP_011614855.1	WP_011614572.1	WP_011614320.1
WP_011615371.1	WP_011615107.1	WP_011614854.1	WP_011614571.1	WP_011614319.1
WP_011615370.1	WP_011615106.1	WP_011614853.1	WP_011614570.1	WP_011614315.1
WP_011615369.1	WP_011615105.1	WP_011614851.1	WP_011614569.1	WP_011614314.1
WP_011615368.1	WP_011615104.1	WP_011614850.1	WP_011614568.1	WP_011614313.1
WP_011615367.1	WP_011615103.1	WP_011614849.1	WP_011614567.1	WP_011614312.1
WP_011615366.1	WP_011615100.1	WP_011614848.1	WP_011614566.1	WP_011614311.1
WP_011615365.1	WP_011615098.1	WP_011614847.1	WP_011614565.1	WP_011614309.1
WP_011615364.1	WP_011615097.1	WP_011614845.1	WP_011614563.1	WP_011614308.1
WP_011615363.1	WP_011615096.1	WP_011614844.1	WP_011614561.1	WP_011614306.1
WP_011615361.1	WP_011615095.1	WP_011614843.1	WP_011614560.1	WP_011614305.1
WP_011615360.1	WP_011615094.1	WP_011614841.1	WP_011614559.1	WP_011614304.1
WP_011615358.1	WP_011615093.1	WP_011614840.1	WP_011614558.1	WP_011614303.1
WP_011615357.1	WP_011615092.1	WP_011614838.1	WP_011614557.1	WP_011614298.1
WP_011615356.1	WP_011615091.1	WP_011614837.1	WP_011614555.1	WP_011614296.1
WP_011615355.1	WP_011615090.1	WP_011614836.1	WP_011614554.1	WP_011614295.1
WP_011615353.1	WP_011615089.1	WP_011614835.1	WP_011614553.1	WP_011614294.1
WP_011615348.1	WP_011615087.1	WP_011614834.1	WP_011614552.1	WP_011614292.1
WP_011615347.1	WP_011615086.1	WP_011614833.1	WP_011614551.1	WP_011614291.1
WP_011615346.1	WP_011615085.1	WP_011614832.1	WP_011614550.1	WP_011614288.1
WP_011615345.1	WP_011615082.1	WP_011614831.1	WP_011614549.1	WP_011614287.1
WP_011615343.1	WP_011615081.1	WP_011614826.1	WP_011614548.1	WP_011614283.1
WP_011615337.1	WP_011615080.1	WP_011614825.1	WP_011614547.1	WP_011614281.1
WP_011615335.1	WP_011615079.1	WP_011614823.1	WP_011614546.1	WP_011614280.1
WP_011615334.1	WP_011615077.1	WP_011614820.1	WP_011614545.1	WP_011614279.1
WP_011615333.1	WP_011615076.1	WP_011614818.1	WP_011614544.1	WP_011614278.1

WP_011614277.1	WP_010814747.1	WP_010814178.1	WP_010813455.1	WP_010812950.1
WP_011614276.1	WP_010814746.1	WP_010814177.1	WP_010813454.1	WP_010812949.1
WP_011614275.1	WP_010814745.1	WP_010814175.1	WP_010813448.1	WP_010812948.1
WP_011153965.1	WP_010814744.1	WP_010814173.1	WP_010813445.1	WP_010812947.1
WP_010815077.1	WP_010814739.1	WP_010814172.1	WP_010813442.1	WP_010812945.1
WP_010815074.1	WP_010814737.1	WP_010813941.1	WP_010813440.1	WP_010812944.1
WP_010815073.1	WP_010814736.1	WP_010813940.1	WP_010813439.1	WP_010812943.1
WP_010815070.1	WP_010814735.1	WP_010813937.1	WP_010813433.1	WP_010812938.1
WP_010815069.1	WP_010814734.1	WP_010813935.1	WP_010813432.1	WP_010812937.1
WP_010815065.1	WP_010814719.1	WP_010813934.1	WP_010813430.1	WP_010812934.1
WP_010815061.1	WP_010814713.1	WP_010813932.1	WP_010813429.1	WP_010812932.1

Appendix B

WP_010815060.1	WP_010814712.1	WP_010813929.1	WP_010813359.1	WP_010812930.1
WP_010815057.1	WP_010814709.1	WP_010813920.1	WP_010813356.1	WP_010812929.1
WP_010815054.1	WP_010814708.1	WP_010813918.1	WP_010813355.1	WP_010812928.1
WP_010815053.1	WP_010814707.1	WP_010813917.1	WP_010813354.1	WP_010812927.1
WP_010815051.1	WP_010814706.1	WP_010813916.1	WP_010813353.1	WP_010812926.1
WP_010815050.1	WP_010814705.1	WP_010813915.1	WP_010813346.1	WP_010812924.1
WP_010815049.1	WP_010814704.1	WP_010813914.1	WP_010813345.1	WP_010812923.1
WP_010815048.1	WP_010814700.1	WP_010813913.1	WP_010813344.1	WP_010812922.1
WP_010815047.1	WP_010814699.1	WP_010813912.1	WP_010813338.1	WP_010812918.1
WP_010815046.1	WP_010814696.1	WP_010813911.1	WP_010813328.1	WP_010812915.1
WP_010815045.1	WP_010814694.1	WP_010813909.1	WP_010813326.1	WP_010812914.1
WP_010815040.1	WP_010814692.1	WP_010813908.1	WP_010813323.1	WP_010812912.1
WP_010815036.1	WP_010814691.1	WP_010813907.1	WP_010813322.1	WP_010812909.1
WP_010815035.1	WP_010814690.1	WP_010813906.1	WP_010813318.1	WP_010812906.1
WP_010815033.1	WP_010814687.1	WP_010813904.1	WP_010813317.1	WP_010812905.1
WP_010815031.1	WP_010814686.1	WP_010813899.1	WP_010813316.1	WP_010812904.1
WP_010815030.1	WP_010814685.1	WP_010813898.1	WP_010813313.1	WP_010812903.1
WP_010815029.1	WP_010814683.1	WP_010813894.1	WP_010813312.1	WP_010812902.1
WP_010815028.1	WP_010814682.1	WP_010813893.1	WP_010813310.1	WP_010812899.1
WP_010815027.1	WP_010814680.1	WP_010813892.1	WP_010813308.1	WP_010812898.1
WP_010815024.1	WP_010814679.1	WP_010813891.1	WP_010813307.1	WP_010812897.1
WP_010815023.1	WP_010814677.1	WP_010813889.1	WP_010813306.1	WP_010812892.1
WP_010815021.1	WP_010814676.1	WP_010813888.1	WP_010813304.1	WP_010812891.1
WP_010815018.1	WP_010814675.1	WP_010813887.1	WP_010813300.1	WP_010812890.1
WP_010815017.1	WP_010814673.1	WP_010813886.1	WP_010813297.1	WP_010812889.1
WP_010815015.1	WP_010814672.1	WP_010813885.1	WP_010813296.1	WP_010812888.1
WP_010815013.1	WP_010814671.1	WP_010813884.1	WP_010813295.1	WP_010812887.1
WP_010815012.1	WP_010814668.1	WP_010813881.1	WP_010813293.1	WP_010812886.1
WP_010815011.1	WP_010814666.1	WP_010813880.1	WP_010813292.1	WP_010812885.1
WP_010815010.1	WP_010814665.1	WP_010813875.1	WP_010813289.1	WP_010812883.1
WP_010815009.1	WP_010814663.1	WP_010813874.1	WP_010813288.1	WP_010812882.1
WP_010815008.1	WP_010814661.1	WP_010813872.1	WP_010813285.1	WP_010812881.1
WP_010815007.1	WP_010814660.1	WP_010813871.1	WP_010813284.1	WP_010812880.1
WP_010815005.1	WP_010814659.1	WP_010813869.1	WP_010813282.1	WP_010812879.1
WP_010815004.1	WP_010814658.1	WP_010813865.1	WP_010813281.1	WP_010812875.1
WP_010815003.1	WP_010814655.1	WP_010813864.1	WP_010813276.1	WP_010812873.1
WP_010815002.1	WP_010814654.1	WP_010813860.1	WP_010813275.1	WP_010812400.1
WP_010815001.1	WP_010814653.1	WP_010813856.1	WP_010813274.1	WP_010812397.1
WP_010814999.1	WP_010814652.1	WP_010813855.1	WP_010813273.1	WP_010812396.1
WP_010814998.1	WP_010814651.1	WP_010813854.1	WP_010813271.1	WP_010812395.1
WP_010814997.1	WP_010814650.1	WP_010813853.1	WP_010813269.1	WP_010812394.1
WP_010814996.1	WP_010814649.1	WP_010813851.1	WP_010813268.1	WP_010812393.1

Appendix B

WP_010814992.1	WP_010814647.1	WP_010813850.1	WP_010813264.1	WP_010812392.1
WP_010814990.1	WP_010814645.1	WP_010813848.1	WP_010813263.1	WP_010812391.1
WP_010814989.1	WP_010814643.1	WP_010813847.1	WP_010813262.1	WP_010812390.1
WP_010814986.1	WP_010814642.1	WP_010813846.1	WP_010813261.1	WP_010812389.1
WP_010814985.1	WP_010814641.1	WP_010813844.1	WP_010813259.1	WP_010812388.1
WP_010814984.1	WP_010814640.1	WP_010813843.1	WP_010813257.1	WP_010812387.1
WP_010814983.1	WP_010814638.1	WP_010813840.1	WP_010813224.1	WP_010812386.1
WP_010814981.1	WP_010814637.1	WP_010813839.1	WP_010813223.1	WP_010812385.1
WP_010814980.1	WP_010814635.1	WP_010813837.1	WP_010813222.1	WP_010812384.1
WP_010814979.1	WP_010814634.1	WP_010813836.1	WP_010813220.1	WP_010812383.1
WP_010814978.1	WP_010814633.1	WP_010813835.1	WP_010813216.1	WP_010812382.1
WP_010814977.1	WP_010814630.1	WP_010813834.1	WP_010813215.1	WP_010812381.1
WP_010814976.1	WP_010814629.1	WP_010813833.1	WP_010813212.1	WP_010812380.1
WP_010814975.1	WP_010814628.1	WP_010813830.1	WP_010813209.1	WP_010812379.1
WP_010814974.1	WP_010814627.1	WP_010813829.1	WP_010813208.1	WP_010812378.1
WP_010814973.1	WP_010814626.1	WP_010813828.1	WP_010813207.1	WP_010812377.1
WP_010814972.1	WP_010814625.1	WP_010813827.1	WP_010813204.1	WP_010812376.1
WP_010814971.1	WP_010814623.1	WP_010813826.1	WP_010813201.1	WP_010812375.1
WP_010814970.1	WP_010814621.1	WP_010813825.1	WP_010813200.1	WP_010812374.1
WP_010814967.1	WP_010814620.1	WP_010813824.1	WP_010813199.1	WP_010812373.1
WP_010814966.1	WP_010814619.1	WP_010813787.1	WP_010813198.1	WP_010812369.1
WP_010814965.1	WP_010814618.1	WP_010813786.1	WP_010813196.1	WP_010812368.1
WP_010814964.1	WP_010814617.1	WP_010813784.1	WP_010813194.1	WP_010812367.1
WP_010814963.1	WP_010814616.1	WP_010813779.1	WP_010813193.1	WP_010812363.1
WP_010814962.1	WP_010814614.1	WP_010813778.1	WP_010813192.1	WP_010812361.1
WP_010814959.1	WP_010814613.1	WP_010813775.1	WP_010813191.1	WP_010812360.1
WP_010814958.1	WP_010814445.1	WP_010813773.1	WP_010813188.1	WP_010812359.1
WP_010814955.1	WP_010814437.1	WP_010813770.1	WP_010813186.1	WP_010812358.1
WP_010814954.1	WP_010814436.1	WP_010813769.1	WP_010813184.1	WP_010812355.1
WP_010814953.1	WP_010814435.1	WP_010813765.1	WP_010813183.1	WP_010812354.1
WP_010814951.1	WP_010814433.1	WP_010813764.1	WP_010813182.1	WP_010812352.1
WP_010814950.1	WP_010814431.1	WP_010813763.1	WP_010813180.1	WP_010812351.1
WP_010814949.1	WP_010814430.1	WP_010813762.1	WP_010813179.1	WP_010812350.1
WP_010814947.1	WP_010814428.1	WP_010813761.1	WP_010813178.1	WP_010812349.1
WP_010814946.1	WP_010814427.1	WP_010813759.1	WP_010813175.1	WP_010812346.1
WP_010814943.1	WP_010814426.1	WP_010813758.1	WP_010813173.1	WP_010812343.1
WP_010814942.1	WP_010814424.1	WP_010813757.1	WP_010813171.1	WP_010812342.1
WP_010814941.1	WP_010814417.1	WP_010813755.1	WP_010813168.1	WP_010812341.1
WP_010814936.1	WP_010814416.1	WP_010813754.1	WP_010813167.1	WP_010812338.1
WP_010814935.1	WP_010814410.1	WP_010813743.1	WP_010813166.1	WP_010812337.1
WP_010814934.1	WP_010814400.1	WP_010813742.1	WP_010813165.1	WP_010812335.1
WP_010814932.1	WP_010814399.1	WP_010813741.1	WP_010813164.1	WP_010812333.1

Appendix B

WP_010814927.1	WP_010814398.1	WP_010813740.1	WP_010813161.1	WP_010812332.1
WP_010814926.1	WP_010814397.1	WP_010813739.1	WP_010813156.1	WP_010812331.1
WP_010814925.1	WP_010814396.1	WP_010813737.1	WP_010813155.1	WP_010812330.1
WP_010814923.1	WP_010814394.1	WP_010813736.1	WP_010813149.1	WP_010812329.1
WP_010814921.1	WP_010814393.1	WP_010813735.1	WP_010813146.1	WP_010812328.1
WP_010814920.1	WP_010814392.1	WP_010813730.1	WP_010813145.1	WP_010812327.1
WP_010814915.1	WP_010814391.1	WP_010813728.1	WP_010813142.1	WP_010812326.1
WP_010814913.1	WP_010814390.1	WP_010813727.1	WP_010813141.1	WP_010812324.1
WP_010814910.1	WP_010814386.1	WP_010813726.1	WP_010813134.1	WP_010812323.1
WP_010814906.1	WP_010814385.1	WP_010813725.1	WP_010813132.1	WP_010812321.1
WP_010814905.1	WP_010814384.1	WP_010813723.1	WP_010813131.1	WP_010812320.1
WP_010814904.1	WP_010814382.1	WP_010813722.1	WP_010813130.1	WP_010812319.1
WP_010814903.1	WP_010814381.1	WP_010813721.1	WP_010813128.1	WP_010812317.1
WP_010814901.1	WP_010814376.1	WP_010813719.1	WP_010813127.1	WP_010812316.1
WP_010814899.1	WP_010814374.1	WP_010813718.1	WP_010813126.1	WP_010812315.1
WP_010814897.1	WP_010814373.1	WP_010813716.1	WP_010813125.1	WP_010812314.1
WP_010814895.1	WP_010814372.1	WP_010813715.1	WP_010813124.1	WP_010812312.1
WP_010814888.1	WP_010814370.1	WP_010813712.1	WP_010813119.1	WP_010812311.1
WP_010814887.1	WP_010814369.1	WP_010813711.1	WP_010813116.1	WP_010812310.1
WP_010814886.1	WP_010814366.1	WP_010813707.1	WP_010813114.1	WP_010812302.1
WP_010814885.1	WP_010814364.1	WP_010813706.1	WP_010813112.1	WP_010812299.1
WP_010814883.1	WP_010814363.1	WP_010813704.1	WP_010813111.1	WP_010812298.1
WP_010814882.1	WP_010814361.1	WP_010813701.1	WP_010813109.1	WP_010812297.1
WP_010814880.1	WP_010814359.1	WP_010813700.1	WP_010813108.1	WP_010812295.1
WP_010814879.1	WP_010814358.1	WP_010813699.1	WP_010813107.1	WP_010812294.1
WP_010814878.1	WP_010814357.1	WP_010813698.1	WP_010813105.1	WP_010812293.1
WP_010814876.1	WP_010814356.1	WP_010813695.1	WP_010813101.1	WP_010812292.1
WP_010814875.1	WP_010814355.1	WP_010813691.1	WP_010813100.1	WP_010812291.1
WP_010814874.1	WP_010814354.1	WP_010813688.1	WP_010813098.1	WP_010812290.1
WP_010814873.1	WP_010814353.1	WP_010813687.1	WP_010813094.1	WP_010812289.1
WP_010814870.1	WP_010814351.1	WP_010813685.1	WP_010813093.1	WP_010812288.1
WP_010814867.1	WP_010814347.1	WP_010813683.1	WP_010813090.1	WP_010812284.1
WP_010814864.1	WP_010814342.1	WP_010813681.1	WP_010813085.1	WP_010812283.1
WP_010814863.1	WP_010814339.1	WP_010813675.1	WP_010813081.1	WP_010812282.1
WP_010814862.1	WP_010814337.1	WP_010813673.1	WP_010813079.1	WP_010812281.1
WP_010814861.1	WP_010814336.1	WP_010813672.1	WP_010813076.1	WP_010812280.1
WP_010814860.1	WP_010814335.1	WP_010813671.1	WP_010813075.1	WP_010812279.1
WP_010814859.1	WP_010814333.1	WP_010813670.1	WP_010813074.1	WP_010812277.1
WP_010814857.1	WP_010814332.1	WP_010813669.1	WP_010813069.1	WP_010812275.1
WP_010814855.1	WP_010814329.1	WP_010813668.1	WP_010813068.1	WP_010812274.1
WP_010814854.1	WP_010814327.1	WP_010813664.1	WP_010813066.1	WP_010812272.1
WP_010814853.1	WP_010814324.1	WP_010813663.1	WP_010813065.1	WP_010812269.1

Appendix B

WP_010814849.1	WP_010814323.1	WP_010813655.1	WP_010813064.1	WP_010812268.1
WP_010814848.1	WP_010814322.1	WP_010813654.1	WP_010813062.1	WP_010812265.1
WP_010814847.1	WP_010814320.1	WP_010813653.1	WP_010813054.1	WP_010812264.1
WP_010814845.1	WP_010814319.1	WP_010813652.1	WP_010813052.1	WP_010812260.1
WP_010814844.1	WP_010814318.1	WP_010813651.1	WP_010813047.1	WP_010812258.1
WP_010814843.1	WP_010814317.1	WP_010813646.1	WP_010813046.1	WP_010812256.1
WP_010814840.1	WP_010814316.1	WP_010813645.1	WP_010813044.1	WP_010812255.1
WP_010814839.1	WP_010814315.1	WP_010813643.1	WP_010813043.1	WP_010812254.1
WP_010814838.1	WP_010814313.1	WP_010813642.1	WP_010813042.1	WP_010812253.1
WP_010814836.1	WP_010814312.1	WP_010813641.1	WP_010813041.1	WP_010812251.1
WP_010814835.1	WP_010814309.1	WP_010813640.1	WP_010813040.1	WP_010812249.1
WP_010814834.1	WP_010814308.1	WP_010813639.1	WP_010813039.1	WP_010812247.1
WP_010814833.1	WP_010814307.1	WP_010813637.1	WP_010813038.1	WP_010812246.1
WP_010814831.1	WP_010814306.1	WP_010813636.1	WP_010813037.1	WP_010812244.1
WP_010814830.1	WP_010814304.1	WP_010813628.1	WP_010813036.1	WP_010812242.1
WP_010814829.1	WP_010814302.1	WP_010813625.1	WP_010813035.1	WP_010812240.1
WP_010814828.1	WP_010814300.1	WP_010813624.1	WP_010813034.1	WP_010812239.1
WP_010814826.1	WP_010814299.1	WP_010813622.1	WP_010813033.1	WP_010812237.1
WP_010814820.1	WP_010814296.1	WP_010813621.1	WP_010813032.1	WP_010812233.1
WP_010814818.1	WP_010814292.1	WP_010813619.1	WP_010813031.1	WP_010812232.1
WP_010814817.1	WP_010814286.1	WP_010813618.1	WP_010813030.1	WP_010812229.1
WP_010814816.1	WP_010814285.1	WP_010813617.1	WP_010813029.1	WP_010812227.1
WP_010814815.1	WP_010814281.1	WP_010813616.1	WP_010813025.1	WP_010812226.1
WP_010814813.1	WP_010814280.1	WP_010813615.1	WP_010813024.1	WP_010812225.1
WP_010814810.1	WP_010814279.1	WP_010813614.1	WP_010813023.1	WP_010812223.1
WP_010814808.1	WP_010814278.1	WP_010813613.1	WP_010813021.1	WP_010812222.1
WP_010814807.1	WP_010814276.1	WP_010813612.1	WP_010813018.1	WP_010812219.1
WP_010814806.1	WP_010814275.1	WP_010813611.1	WP_010813017.1	WP_010812209.1
WP_010814805.1	WP_010814273.1	WP_010813610.1	WP_010813016.1	WP_010812207.1
WP_010814804.1	WP_010814268.1	WP_010813609.1	WP_010813015.1	WP_010812206.1
WP_010814801.1	WP_010814267.1	WP_010813608.1	WP_010813009.1	WP_010812205.1
WP_010814799.1	WP_010814266.1	WP_010813607.1	WP_010813008.1	WP_010812204.1
WP_010814798.1	WP_010814265.1	WP_010813606.1	WP_010813007.1	WP_010812198.1
WP_010814796.1	WP_010814261.1	WP_010813605.1	WP_010813006.1	WP_010812197.1
WP_010814794.1	WP_010814257.1	WP_010813578.1	WP_010813005.1	WP_010812196.1
WP_010814793.1	WP_010814256.1	WP_010813577.1	WP_010813003.1	WP_010812195.1
WP_010814792.1	WP_010814255.1	WP_010813574.1	WP_010813001.1	WP_010812193.1
WP_010814791.1	WP_010814254.1	WP_010813568.1	WP_010813000.1	WP_010812189.1
WP_010814789.1	WP_010814252.1	WP_010813567.1	WP_010812998.1	WP_010812186.1
WP_010814785.1	WP_010814251.1	WP_010813566.1	WP_010812997.1	WP_010812185.1
WP_010814784.1	WP_010814250.1	WP_010813565.1	WP_010812995.1	WP_010812183.1
WP_010814782.1	WP_010814249.1	WP_010813562.1	WP_010812994.1	WP_010812182.1

Appendix B

WP_010814781.1	WP_010814248.1	WP_010813560.1	WP_010812992.1	WP_010812179.1
WP_010814780.1	WP_010814246.1	WP_010813559.1	WP_010812991.1	WP_010812177.1
WP_010814779.1	WP_010814245.1	WP_010813556.1	WP_010812984.1	WP_010812176.1
WP_010814778.1	WP_010814215.1	WP_010813554.1	WP_010812983.1	WP_010812175.1
WP_010814777.1	WP_010814213.1	WP_010813553.1	WP_010812979.1	WP_010812173.1
WP_010814775.1	WP_010814211.1	WP_010813547.1	WP_010812975.1	WP_010812171.1
WP_010814772.1	WP_010814209.1	WP_010813542.1	WP_010812974.1	WP_010812169.1
WP_010814770.1	WP_010814206.1	WP_010813540.1	WP_010812971.1	WP_010812167.1
WP_010814769.1	WP_010814205.1	WP_010813539.1	WP_010812968.1	WP_010812166.1
WP_010814768.1	WP_010814204.1	WP_010813538.1	WP_010812967.1	WP_010812161.1
WP_010814766.1	WP_010814202.1	WP_010813536.1	WP_010812966.1	WP_010812160.1
WP_010814765.1	WP_010814200.1	WP_010813475.1	WP_010812965.1	WP_010812159.1
WP_010814764.1	WP_010814199.1	WP_010813472.1	WP_010812964.1	WP_010812157.1
WP_010814763.1	WP_010814195.1	WP_010813470.1	WP_010812962.1	WP_010812154.1
WP_010814762.1	WP_010814193.1	WP_010813469.1	WP_010812959.1	WP_010812152.1
WP_010814760.1	WP_010814192.1	WP_010813465.1	WP_010812957.1	WP_010812149.1
WP_010814758.1	WP_010814191.1	WP_010813464.1	WP_010812955.1	WP_010812148.1
WP_010814756.1	WP_010814190.1	WP_010813462.1	WP_010812954.1	WP_010812147.1
WP_010814755.1	WP_010814186.1	WP_010813461.1	WP_010812953.1	WP_010812145.1
WP_010814753.1	WP_010814181.1	WP_010813460.1	WP_010812952.1	WP_010812144.1
WP_010814748.1	WP_010814179.1	WP_010813458.1	WP_010812951.1	WP_010812142.1

WP_010812140.1	WP_010811267.1	WP_010809948.1	WP_010809120.1
WP_010812139.1	WP_010811266.1	WP_010809946.1	WP_010809119.1
WP_010812137.1	WP_010811265.1	WP_010809945.1	WP_010809118.1
WP_010812135.1	WP_010811264.1	WP_010809944.1	WP_010809117.1
WP_010812134.1	WP_010811263.1	WP_010809940.1	WP_010809115.1
WP_010812133.1	WP_010811262.1	WP_010809938.1	WP_010809112.1
WP_010812129.1	WP_010811261.1	WP_010809914.1	WP_010809111.1
WP_010812128.1	WP_010811259.1	WP_010809912.1	WP_010809110.1
WP_010812127.1	WP_010811256.1	WP_010809911.1	WP_010809109.1
WP_010812126.1	WP_010811254.1	WP_010809909.1	WP_010809108.1
WP_010812125.1	WP_010811246.1	WP_010809907.1	WP_010809107.1
WP_010812123.1	WP_010811245.1	WP_010809905.1	WP_010809104.1
WP_010812119.1	WP_010811241.1	WP_010809904.1	WP_010809103.1
WP_010812118.1	WP_010811240.1	WP_010809903.1	WP_010809101.1
WP_010812117.1	WP_010811239.1	WP_010809902.1	WP_010809100.1
WP_010812116.1	WP_010811235.1	WP_010809901.1	WP_010809095.1
WP_010812115.1	WP_010811233.1	WP_010809900.1	WP_010809092.1
WP_010812107.1	WP_010811224.1	WP_010809899.1	WP_010809091.1
WP_010812106.1	WP_010811221.1	WP_010809896.1	WP_010809090.1

Appendix B

WP_010812105.1	WP_010811218.1	WP_010809892.1	WP_010809084.1
WP_010812104.1	WP_010811215.1	WP_010809891.1	WP_010809083.1
WP_010812102.1	WP_010810635.1	WP_010809888.1	WP_010809082.1
WP_010812101.1	WP_010810633.1	WP_010809885.1	WP_010809080.1
WP_010812098.1	WP_010810632.1	WP_010809884.1	WP_010809078.1
WP_010812094.1	WP_010810631.1	WP_010809883.1	WP_010809077.1
WP_010812093.1	WP_010810625.1	WP_010809882.1	WP_010809074.1
WP_010812090.1	WP_010810624.1	WP_010809880.1	WP_010809073.1
WP_010812089.1	WP_010810622.1	WP_010809877.1	WP_010809072.1
WP_010812088.1	WP_010810620.1	WP_010809876.1	WP_010809065.1
WP_010812087.1	WP_010810615.1	WP_010809874.1	WP_010809064.1
WP_010812079.1	WP_010810612.1	WP_010809873.1	WP_010809063.1
WP_010812078.1	WP_010810611.1	WP_010809870.1	WP_010809062.1
WP_010812071.1	WP_010810609.1	WP_010809868.1	WP_010809060.1
WP_010812070.1	WP_010810607.1	WP_010809867.1	WP_010809059.1
WP_010812065.1	WP_010810605.1	WP_010809866.1	WP_010809051.1
WP_010812064.1	WP_010810603.1	WP_010809602.1	WP_010809050.1
WP_010812061.1	WP_010810601.1	WP_010809601.1	WP_010809049.1
WP_010812059.1	WP_010810593.1	WP_010809600.1	WP_010809047.1
WP_010812057.1	WP_010810590.1	WP_010809599.1	WP_010809044.1
WP_010812054.1	WP_010810583.1	WP_010809598.1	WP_010809042.1
WP_010812049.1	WP_010810582.1	WP_010809596.1	WP_010809038.1
WP_010812047.1	WP_010810581.1	WP_010809595.1	WP_010809035.1
WP_010812042.1	WP_010810580.1	WP_010809594.1	WP_010809033.1
WP_010812031.1	WP_010810575.1	WP_010809591.1	WP_010809032.1
WP_010812025.1	WP_010810566.1	WP_010809590.1	WP_010809031.1
WP_010812023.1	WP_010810562.1	WP_010809589.1	WP_010809030.1
WP_010812021.1	WP_010810561.1	WP_010809588.1	WP_010809029.1
WP_010812012.1	WP_010810558.1	WP_010809585.1	WP_010809026.1
WP_010812008.1	WP_010810554.1	WP_010809584.1	WP_010809025.1
WP_010811799.1	WP_010810552.1	WP_010809583.1	WP_010809024.1
WP_010811797.1	WP_010810551.1	WP_010809582.1	WP_010809022.1
WP_010811794.1	WP_010810550.1	WP_010809581.1	WP_010809021.1
WP_010811787.1	WP_010810549.1	WP_010809580.1	WP_010809020.1
WP_010811785.1	WP_010810548.1	WP_010809579.1	WP_010809019.1
WP_010811778.1	WP_010810547.1	WP_010809576.1	WP_010809018.1
WP_010811772.1	WP_010810546.1	WP_010809575.1	WP_010809016.1
WP_010811769.1	WP_010810545.1	WP_010809554.1	WP_010809014.1
WP_010811767.1	WP_010810544.1	WP_010809543.1	WP_010809013.1
WP_010811762.1	WP_010810543.1	WP_010809540.1	WP_010809011.1
WP_010811761.1	WP_010810542.1	WP_010809539.1	WP_010809009.1
WP_010811759.1	WP_010810541.1	WP_010809538.1	WP_010809005.1

Appendix B

WP_010811758.1	WP_010810540.1	WP_010809536.1	WP_010809004.1
WP_010811754.1	WP_010810537.1	WP_010809535.1	WP_010809003.1
WP_010811746.1	WP_010810535.1	WP_010809534.1	WP_010809002.1
WP_010811742.1	WP_010810534.1	WP_010809532.1	WP_010809001.1
WP_010811741.1	WP_010810531.1	WP_010809531.1	WP_010809000.1
WP_010811738.1	WP_010810529.1	WP_010809529.1	WP_010808999.1
WP_010811736.1	WP_010810528.1	WP_010809523.1	WP_010808997.1
WP_010811734.1	WP_010810526.1	WP_010809519.1	WP_010808996.1
WP_010811733.1	WP_010810524.1	WP_010809518.1	WP_010808995.1
WP_010811730.1	WP_010810523.1	WP_010809517.1	WP_010808992.1
WP_010811729.1	WP_010810522.1	WP_010809516.1	WP_010808985.1
WP_010811728.1	WP_010810514.1	WP_010809515.1	WP_010808983.1
WP_010811619.1	WP_010810508.1	WP_010809513.1	WP_010808982.1
WP_010811613.1	WP_010810506.1	WP_010809510.1	WP_010808979.1
WP_010811612.1	WP_010810505.1	WP_010809503.1	WP_010808978.1
WP_010811609.1	WP_010810504.1	WP_010809499.1	WP_010808975.1
WP_010811608.1	WP_010810503.1	WP_010809497.1	WP_010808972.1
WP_010811603.1	WP_010810501.1	WP_010809495.1	WP_010808971.1
WP_010811601.1	WP_010810495.1	WP_010809494.1	WP_010808970.1
WP_010811598.1	WP_010810494.1	WP_010809489.1	WP_010808968.1
WP_010811593.1	WP_010810493.1	WP_010809487.1	WP_010808966.1
WP_010811591.1	WP_010810492.1	WP_010809485.1	WP_010808965.1
WP_010811590.1	WP_010810488.1	WP_010809484.1	WP_010808962.1
WP_010811589.1	WP_010810487.1	WP_010809483.1	WP_010808961.1
WP_010811582.1	WP_010810484.1	WP_010809482.1	WP_010808959.1
WP_010811581.1	WP_010810482.1	WP_010809481.1	WP_010808953.1
WP_010811579.1	WP_010810480.1	WP_010809479.1	WP_010808952.1
WP_010811578.1	WP_010810474.1	WP_010809477.1	WP_010808951.1
WP_010811575.1	WP_010810473.1	WP_010809476.1	WP_010808947.1
WP_010811555.1	WP_010810472.1	WP_010809472.1	WP_010808945.1
WP_010811552.1	WP_010810471.1	WP_010809471.1	WP_010808944.1
WP_010811546.1	WP_010810470.1	WP_010809470.1	WP_010808940.1
WP_010811545.1	WP_010810469.1	WP_010809469.1	WP_010808933.1
WP_010811543.1	WP_010810468.1	WP_010809466.1	WP_010808932.1
WP_010811542.1	WP_010810467.1	WP_010809465.1	WP_010808930.1
WP_010811537.1	WP_010810466.1	WP_010809464.1	WP_010808929.1
WP_010811536.1	WP_010810465.1	WP_010809463.1	WP_010808926.1
WP_010811535.1	WP_010810464.1	WP_010809462.1	WP_010808924.1
WP_010811534.1	WP_010810463.1	WP_010809461.1	WP_010808919.1
WP_010811530.1	WP_010810462.1	WP_010809460.1	WP_010808918.1
WP_010811527.1	WP_010810461.1	WP_010809459.1	WP_010808916.1
WP_010811526.1	WP_010810460.1	WP_010809456.1	WP_010808914.1

Appendix B

WP_010811523.1	WP_010810459.1	WP_010809454.1	WP_010808913.1
WP_010811519.1	WP_010810458.1	WP_010809453.1	WP_010808912.1
WP_010811517.1	WP_010810246.1	WP_010809451.1	WP_010808911.1
WP_010811511.1	WP_010810142.1	WP_010809446.1	WP_010808910.1
WP_010811509.1	WP_010810139.1	WP_010809445.1	WP_010808909.1
WP_010811506.1	WP_010810138.1	WP_010809443.1	WP_010808908.1
WP_010811505.1	WP_010810134.1	WP_010809442.1	WP_008650850.1
WP_010811504.1	WP_010810132.1	WP_010809441.1	WP_008650487.1
WP_010811503.1	WP_010810131.1	WP_010809440.1	WP_008648066.1
WP_010811499.1	WP_010810130.1	WP_010809439.1	WP_008644813.1
WP_010811495.1	WP_010810129.1	WP_010809438.1	WP_008643342.1
WP_010811494.1	WP_010810124.1	WP_010809437.1	WP_008642959.1
WP_010811493.1	WP_010810122.1	WP_010809436.1	WP_008641674.1
WP_010811490.1	WP_010810118.1	WP_010809435.1	WP_006578656.1
WP_010811489.1	WP_010810116.1	WP_010809434.1	WP_006576525.1
WP_010811487.1	WP_010810113.1	WP_010809433.1	WP_006576245.1
WP_010811486.1	WP_010810112.1	WP_010809432.1	WP_006575661.1
WP_010811485.1	WP_010810110.1	WP_010809430.1	WP_006575466.1
WP_010811484.1	WP_010810109.1	WP_010809428.1	WP_006163606.1
WP_010811483.1	WP_010810107.1	WP_010809427.1	WP_006162945.1
WP_010811481.1	WP_010810106.1	WP_010809426.1	WP_006162803.1
WP_010811480.1	WP_010810101.1	WP_010809425.1	WP_006160488.1
WP_010811478.1	WP_010810100.1	WP_010809424.1	WP_006160449.1
WP_010811477.1	WP_010810098.1	WP_010809422.1	
WP_010811476.1	WP_010810097.1	WP_010809421.1	
WP_010811473.1	WP_010810096.1	WP_010809419.1	
WP_010811472.1	WP_010810095.1	WP_010809246.1	
WP_010811467.1	WP_010810092.1	WP_010809245.1	
WP_010811462.1	WP_010810087.1	WP_010809244.1	
WP_010811461.1	WP_010810086.1	WP_010809242.1	
WP_010811457.1	WP_010810084.1	WP_010809241.1	
WP_010811456.1	WP_010810072.1	WP_010809240.1	
WP_010811455.1	WP_010810067.1	WP_010809235.1	
WP_010811453.1	WP_010810065.1	WP_010809233.1	
WP_010811452.1	WP_010810064.1	WP_010809232.1	
WP_010811451.1	WP_010810063.1	WP_010809230.1	
WP_010811450.1	WP_010810059.1	WP_010809229.1	
WP_010811449.1	WP_010810058.1	WP_010809226.1	
WP_010811448.1	WP_010810057.1	WP_010809225.1	
WP_010811446.1	WP_010810056.1	WP_010809223.1	
WP_010811444.1	WP_010810055.1	WP_010809222.1	
WP_010811437.1	WP_010810052.1	WP_010809220.1	

Appendix B

WP_010811431.1	WP_010810051.1	WP_010809219.1	
WP_010811430.1	WP_010810050.1	WP_010809217.1	
WP_010811429.1	WP_010810049.1	WP_010809215.1	
WP_010811426.1	WP_010810046.1	WP_010809214.1	
WP_010811418.1	WP_010810044.1	WP_010809211.1	
WP_010811417.1	WP_010810043.1	WP_010809210.1	
WP_010811416.1	WP_010810041.1	WP_010809209.1	
WP_010811414.1	WP_010810040.1	WP_010809205.1	
WP_010811413.1	WP_010810038.1	WP_010809204.1	
WP_010811412.1	WP_010810037.1	WP_010809200.1	
WP_010811407.1	WP_010810036.1	WP_010809195.1	
WP_010811404.1	WP_010810035.1	WP_010809194.1	
WP_010811402.1	WP_010810031.1	WP_010809193.1	
WP_010811401.1	WP_010810028.1	WP_010809187.1	
WP_010811375.1	WP_010810026.1	WP_010809186.1	
WP_010811372.1	WP_010810025.1	WP_010809184.1	
WP_010811358.1	WP_010810022.1	WP_010809179.1	
WP_010811357.1	WP_010810020.1	WP_010809178.1	
WP_010811354.1	WP_010810019.1	WP_010809176.1	
WP_010811352.1	WP_010810016.1	WP_010809175.1	
WP_010811347.1	WP_010810014.1	WP_010809173.1	
WP_010811346.1	WP_010810013.1	WP_010809171.1	
WP_010811345.1	WP_010810012.1	WP_010809170.1	
WP_010811344.1	WP_010810009.1	WP_010809169.1	
WP_010811341.1	WP_010810008.1	WP_010809168.1	
WP_010811338.1	WP_010810007.1	WP_010809166.1	
WP_010811336.1	WP_010810006.1	WP_010809162.1	
WP_010811335.1	WP_010810003.1	WP_010809161.1	
WP_010811334.1	WP_010810002.1	WP_010809160.1	
WP_010811333.1	WP_010810001.1	WP_010809159.1	
WP_010811332.1	WP_010809996.1	WP_010809157.1	
WP_010811325.1	WP_010809995.1	WP_010809156.1	
WP_010811323.1	WP_010809994.1	WP_010809154.1	
WP_010811313.1	WP_010809993.1	WP_010809153.1	
WP_010811311.1	WP_010809983.1	WP_010809150.1	
WP_010811309.1	WP_010809982.1	WP_010809146.1	
WP_010811308.1	WP_010809981.1	WP_010809144.1	
WP_010811304.1	WP_010809980.1	WP_010809143.1	
WP_010811303.1	WP_010809979.1	WP_010809142.1	
WP_010811296.1	WP_010809978.1	WP_010809141.1	
WP_010811293.1	WP_010809972.1	WP_010809140.1	
WP_010811292.1	WP_010809970.1	WP_010809139.1	

WP_010811289.1	WP_010809969.1	WP_010809138.1	
WP_010811288.1	WP_010809968.1	WP_010809137.1	
WP_010811286.1	WP_010809967.1	WP_010809136.1	
WP_010811285.1	WP_010809966.1	WP_010809132.1	
WP_010811281.1	WP_010809963.1	WP_010809131.1	
WP_010811280.1	WP_010809962.1	WP_010809130.1	
WP_010811279.1	WP_010809961.1	WP_010809129.1	
WP_010811276.1	WP_010809959.1	WP_010809128.1	
WP_010811274.1	WP_010809958.1	WP_010809127.1	
WP_010811272.1	WP_010809956.1	WP_010809125.1	
WP_010811271.1	WP_010809955.1	WP_010809124.1	
WP_010811270.1	WP_010809954.1	WP_010809123.1	
WP_010811268.1	WP_010809949.1	WP_010809122.1	

*The green cell presented the type 1 BVMO encoding regions.

Table 43. The nucleotide encoding regions in chromosome 2 of *C.necator* H16.

WP_197540027.1	WP_041688442.1	WP_011617949.1	WP_011617684.1	WP_011617425.1
WP_197540026.1	WP_041688428.1	WP_011617946.1	WP_011617682.1	WP_011617424.1
WP_197540025.1	WP_041688413.1	WP_011617945.1	WP_011617680.1	WP_011617423.1
WP_197540024.1	WP_041688404.1	WP_011617943.1	WP_011617679.1	WP_011617422.1
WP_197540023.1	WP_041688394.1	WP_011617942.1	WP_011617678.1	WP_011617421.1
WP_197540022.1	WP_041688388.1	WP_011617941.1	WP_011617677.1	WP_011617419.1
WP_197540021.1	WP_041688386.1	WP_011617940.1	WP_011617676.1	WP_011617418.1
WP_193385850.1	WP_041688376.1	WP_011617939.1	WP_011617675.1	WP_011617417.1
WP_193385849.1	WP_041688372.1	WP_011617938.1	WP_011617672.1	WP_011617415.1
WP_174549481.1	WP_041688347.1	WP_011617937.1	WP_011617670.1	WP_011617414.1
WP_174549474.1	WP_041688341.1	WP_011617936.1	WP_011617669.1	WP_011617413.1
WP_174549473.1	WP_041688337.1	WP_011617935.1	WP_011617668.1	WP_011617412.1
WP_174549464.1	WP_041688333.1	WP_011617934.1	WP_011617667.1	WP_011617411.1
WP_174549463.1	WP_041688331.1	WP_011617932.1	WP_011617665.1	WP_011617410.1
WP_174549462.1	WP_041688319.1	WP_011617931.1	WP_011617664.1	WP_011617408.1
WP_174549461.1	WP_041688313.1	WP_011617930.1	WP_011617663.1	WP_011617407.1
WP_174549459.1	WP_041688295.1	WP_011617929.1	WP_011617662.1	WP_011617406.1
WP_174549458.1	WP_041688276.1	WP_011617928.1	WP_011617661.1	WP_011617405.1
WP_174549457.1	WP_041688275.1	WP_011617927.1	WP_011617660.1	WP_011617404.1
WP_174549454.1	WP_041688273.1	WP_011617926.1	WP_011617659.1	WP_011617403.1
WP_174549449.1	WP_041688258.1	WP_011617925.1	WP_011617658.1	WP_011617402.1
WP_174549448.1	WP_041688256.1	WP_011617924.1	WP_011617657.1	WP_011617401.1
WP_174549446.1	WP_041688254.1	WP_011617923.1	WP_011617656.1	WP_011617397.1
WP_174549445.1	WP_041688251.1	WP_011617922.1	WP_011617655.1	WP_011617396.1
WP_167686603.1	WP_041688247.1	WP_011617921.1	WP_011617654.1	WP_011617395.1
WP_167686602.1	WP_041688245.1	WP_011617920.1	WP_011617653.1	WP_011617393.1

Appendix B

WP_167686601.1	WP_041688242.1	WP_011617919.1	WP_011617652.1	WP_011617392.1
WP_167686600.1	WP_041688240.1	WP_011617918.1	WP_011617650.1	WP_011617391.1
WP_167686599.1	WP_041688238.1	WP_011617917.1	WP_011617649.1	WP_011617390.1
WP_167686598.1	WP_041688234.1	WP_011617916.1	WP_011617648.1	WP_011617387.1
WP_167686597.1	WP_041688232.1	WP_011617915.1	WP_011617645.1	WP_011617386.1
WP_167686596.1	WP_041688227.1	WP_011617914.1	WP_011617644.1	WP_011617385.1
WP_167686595.1	WP_041688226.1	WP_011617913.1	WP_011617643.1	WP_011617382.1
WP_167686594.1	WP_041688225.1	WP_011617912.1	WP_011617642.1	WP_011617380.1
WP_167686593.1	WP_041688214.1	WP_011617911.1	WP_011617641.1	WP_011617379.1
WP_167686592.1	WP_041688212.1	WP_011617909.1	WP_011617640.1	WP_011617378.1
WP_167686591.1	WP_041688210.1	WP_011617908.1	WP_011617639.1	WP_011617377.1
WP_167686590.1	WP_041688208.1	WP_011617905.1	WP_011617638.1	WP_011617376.1
WP_162166399.1	WP_041688204.1	WP_011617904.1	WP_011617636.1	WP_011617375.1
WP_161940223.1	WP_041688202.1	WP_011617903.1	WP_011617634.1	WP_011617374.1
WP_157886312.1	WP_041688198.1	WP_011617901.1	WP_011617633.1	WP_011617373.1
WP_157886310.1	WP_041688196.1	WP_011617900.1	WP_011617631.1	WP_011617371.1
WP_157886309.1	WP_041688194.1	WP_011617899.1	WP_011617628.1	WP_011617370.1
WP_157886307.1	WP_041688192.1	WP_011617898.1	WP_011617627.1	WP_011617369.1
WP_157886305.1	WP_041688190.1	WP_011617897.1	WP_011617626.1	WP_011617368.1
WP_155737100.1	WP_041688188.1	WP_011617896.1	WP_011617625.1	WP_011617367.1
WP_155737077.1	WP_041688181.1	WP_011617895.1	WP_011617624.1	WP_011617365.1
WP_155737074.1	WP_041688178.1	WP_011617894.1	WP_011617623.1	WP_011617364.1
WP_155737070.1	WP_041688176.1	WP_011617892.1	WP_011617622.1	WP_011617363.1
WP_155737069.1	WP_041688172.1	WP_011617891.1	WP_011617621.1	WP_011617362.1
WP_136227926.1	WP_041688170.1	WP_011617890.1	WP_011617620.1	WP_011617361.1
WP_136227925.1	WP_041688168.1	WP_011617888.1	WP_011617619.1	WP_011617360.1
WP_136227923.1	WP_041688165.1	WP_011617887.1	WP_011617618.1	WP_011617359.1
WP_136227922.1	WP_041688163.1	WP_011617886.1	WP_011617617.1	WP_011617358.1
WP_136227921.1	WP_041688159.1	WP_011617881.1	WP_011617616.1	WP_011617357.1
WP_136227920.1	WP_041688156.1	WP_011617880.1	WP_011617614.1	WP_011617356.1
WP_136227919.1	WP_041688154.1	WP_011617879.1	WP_011617613.1	WP_011617355.1
WP_136227918.1	WP_041688152.1	WP_011617878.1	WP_011617612.1	WP_011617354.1
WP_136227917.1	WP_041688146.1	WP_011617877.1	WP_011617611.1	WP_011617353.1
WP_136227916.1	WP_041688140.1	WP_011617873.1	WP_011617610.1	WP_011617352.1
WP_136227915.1	WP_041688138.1	WP_011617872.1	WP_011617609.1	WP_011617350.1
WP_136227914.1	WP_041688134.1	WP_011617871.1	WP_011617607.1	WP_011617349.1
WP_136227913.1	WP_041688132.1	WP_011617870.1	WP_011617606.1	WP_011617347.1
WP_136227912.1	WP_041688130.1	WP_011617869.1	WP_011617605.1	WP_011617346.1
WP_136227911.1	WP_041688127.1	WP_011617867.1	WP_011617600.1	WP_011617345.1
WP_136227909.1	WP_041688123.1	WP_011617866.1	WP_011617595.1	WP_011617344.1
WP_136227908.1	WP_041688121.1	WP_011617865.1	WP_011617593.1	WP_011617343.1
WP_136227906.1	WP_041688119.1	WP_011617864.1	WP_011617592.1	WP_011617342.1

Appendix B

WP_136227905.1	WP_041688113.1	WP_011617863.1	WP_011617591.1	WP_011617341.1
WP_136227904.1	WP_041688111.1	WP_011617862.1	WP_011617589.1	WP_011617339.1
WP_136227903.1	WP_041688108.1	WP_011617861.1	WP_011617587.1	WP_011617338.1
WP_136227902.1	WP_041688106.1	WP_011617860.1	WP_011617586.1	WP_011617337.1
WP_136227901.1	WP_041688104.1	WP_011617857.1	WP_011617585.1	WP_011617336.1
WP_136227900.1	WP_041688100.1	WP_011617856.1	WP_011617584.1	WP_011617335.1
WP_136227899.1	WP_041688097.1	WP_011617854.1	WP_011617583.1	WP_011617334.1
WP_136227898.1	WP_041688095.1	WP_011617853.1	WP_011617581.1	WP_011617333.1
WP_136227897.1	WP_041688089.1	WP_011617852.1	WP_011617580.1	WP_011617332.1
WP_136227896.1	WP_041688087.1	WP_011617851.1	WP_011617579.1	WP_011617331.1
WP_136227895.1	WP_041688085.1	WP_011617850.1	WP_011617578.1	WP_011617329.1
WP_136227894.1	WP_041688083.1	WP_011617849.1	WP_011617576.1	WP_011617328.1
WP_136227893.1	WP_041688081.1	WP_011617848.1	WP_011617575.1	WP_011617326.1
WP_136227892.1	WP_041688070.1	WP_011617847.1	WP_011617573.1	WP_011617325.1
WP_136227889.1	WP_041688068.1	WP_011617846.1	WP_011617572.1	WP_011617324.1
WP_136227888.1	WP_041688066.1	WP_011617844.1	WP_011617571.1	WP_011617323.1
WP_136227886.1	WP_041688064.1	WP_011617842.1	WP_011617567.1	WP_011617322.1
WP_136227885.1	WP_041688061.1	WP_011617841.1	WP_011617566.1	WP_011617321.1
WP_136227884.1	WP_041688058.1	WP_011617839.1	WP_011617565.1	WP_011617320.1
WP_136227883.1	WP_041688054.1	WP_011617835.1	WP_011617564.1	WP_011617319.1
WP_136227881.1	WP_041688052.1	WP_011617833.1	WP_011617563.1	WP_011617318.1
WP_136227880.1	WP_041688050.1	WP_011617832.1	WP_011617562.1	WP_011617316.1
WP_136227878.1	WP_041688048.1	WP_011617831.1	WP_011617561.1	WP_011617314.1
WP_136227877.1	WP_041688046.1	WP_011617830.1	WP_011617560.1	WP_011617312.1
WP_136227876.1	WP_041688042.1	WP_011617829.1	WP_011617558.1	WP_011617311.1
WP_136227875.1	WP_041688040.1	WP_011617828.1	WP_011617556.1	WP_011617310.1
WP_136227874.1	WP_041688036.1	WP_011617827.1	WP_011617554.1	WP_011617309.1
WP_136227872.1	WP_041688032.1	WP_011617825.1	WP_011617553.1	WP_011617308.1
WP_136227871.1	WP_041688028.1	WP_011617823.1	WP_011617552.1	WP_011617307.1
WP_136227868.1	WP_041688026.1	WP_011617821.1	WP_011617551.1	WP_011617306.1
WP_136227866.1	WP_041688024.1	WP_011617820.1	WP_011617550.1	WP_011617305.1
WP_136227865.1	WP_041688022.1	WP_011617819.1	WP_011617549.1	WP_011617304.1
WP_136227862.1	WP_041688019.1	WP_011617817.1	WP_011617548.1	WP_011617303.1
WP_136227861.1	WP_041688015.1	WP_011617816.1	WP_011617546.1	WP_011617301.1
WP_136227860.1	WP_041688010.1	WP_011617815.1	WP_011617545.1	WP_011617300.1
WP_136227859.1	WP_041688004.1	WP_011617814.1	WP_011617544.1	WP_011617299.1
WP_136227858.1	WP_041688002.1	WP_011617813.1	WP_011617543.1	WP_011617297.1
WP_136227857.1	WP_041687997.1	WP_011617812.1	WP_011617542.1	WP_011617296.1
WP_136227853.1	WP_041687995.1	WP_011617808.1	WP_011617541.1	WP_011617295.1
WP_136227852.1	WP_041687992.1	WP_011617807.1	WP_011617540.1	WP_011617294.1
WP_136227851.1	WP_041687986.1	WP_011617806.1	WP_011617539.1	WP_011617293.1
WP_136227849.1	WP_041687979.1	WP_011617804.1	WP_011617538.1	WP_011617292.1

Appendix B

WP_136227848.1	WP_041687977.1	WP_011617803.1	WP_011617536.1	WP_011617291.1
WP_136227847.1	WP_041687975.1	WP_011617802.1	WP_011617535.1	WP_011617290.1
WP_136227844.1	WP_041687973.1	WP_011617801.1	WP_011617534.1	WP_011617289.1
WP_136227843.1	WP_041687970.1	WP_011617800.1	WP_011617532.1	WP_011617288.1
WP_136227841.1	WP_041687964.1	WP_011617799.1	WP_011617531.1	WP_011617287.1
WP_136227840.1	WP_041687954.1	WP_011617798.1	WP_011617530.1	WP_011617286.1
WP_136227839.1	WP_041687953.1	WP_011617797.1	WP_011617529.1	WP_011617284.1
WP_136227837.1	WP_041687951.1	WP_011617796.1	WP_011617528.1	WP_011617283.1
WP_136227836.1	WP_041687949.1	WP_011617795.1	WP_011617527.1	WP_011617282.1
WP_099045631.1	WP_041687941.1	WP_011617794.1	WP_011617526.1	WP_011617281.1
WP_099045606.1	WP_041687935.1	WP_011617792.1	WP_011617524.1	WP_011617280.1
WP_082236113.1	WP_041687930.1	WP_011617791.1	WP_011617523.1	WP_011617278.1
WP_081225831.1	WP_041687928.1	WP_011617790.1	WP_011617522.1	WP_011617276.1
WP_081225830.1	WP_041687924.1	WP_011617789.1	WP_011617521.1	WP_011617275.1
WP_081225829.1	WP_041687922.1	WP_011617788.1	WP_011617520.1	WP_011617273.1
WP_081225826.1	WP_041687915.1	WP_011617787.1	WP_011617519.1	WP_011617271.1
WP_081225824.1	WP_041687914.1	WP_011617786.1	WP_011617518.1	WP_011617269.1
WP_081225814.1	WP_041687912.1	WP_011617783.1	WP_011617517.1	WP_011617268.1
WP_081225793.1	WP_041687909.1	WP_011617780.1	WP_011617516.1	WP_011617267.1
WP_081225792.1	WP_041687906.1	WP_011617777.1	WP_011617515.1	WP_011617266.1
WP_081225791.1	WP_041687904.1	WP_011617776.1	WP_011617514.1	WP_011617265.1
WP_081225786.1	WP_041687902.1	WP_011617775.1	WP_011617513.1	WP_011617264.1
WP_081225781.1	WP_041687895.1	WP_011617774.1	WP_011617512.1	WP_011617263.1
WP_081225780.1	WP_041687890.1	WP_011617773.1	WP_011617510.1	WP_011617262.1
WP_081225779.1	WP_039016261.1	WP_011617771.1	WP_011617509.1	WP_011617261.1
WP_081225778.1	WP_037025844.1	WP_011617767.1	WP_011617508.1	WP_011617260.1
WP_081225766.1	WP_037025841.1	WP_011617765.1	WP_011617507.1	WP_011617259.1
WP_081225761.1	WP_037025720.1	WP_011617764.1	WP_011617506.1	WP_011617258.1
WP_081225752.1	WP_037025644.1	WP_011617763.1	WP_011617505.1	WP_011617257.1
WP_081225751.1	WP_037025585.1	WP_011617762.1	WP_011617504.1	WP_011617255.1
WP_081225750.1	WP_037025578.1	WP_011617761.1	WP_011617503.1	WP_011617254.1
WP_081225749.1	WP_037025558.1	WP_011617760.1	WP_011617502.1	WP_011617253.1
WP_081225744.1	WP_037025549.1	WP_011617759.1	WP_011617499.1	WP_011617250.1
WP_081050364.1	WP_037025516.1	WP_011617758.1	WP_011617498.1	WP_011617249.1
WP_081050325.1	WP_037025481.1	WP_011617757.1	WP_011617497.1	WP_011617248.1
WP_081050322.1	WP_037025473.1	WP_011617756.1	WP_011617496.1	WP_011617244.1
WP_081050256.1	WP_037025456.1	WP_011617755.1	WP_011617495.1	WP_011617243.1
WP_081050246.1	WP_037025453.1	WP_011617754.1	WP_011617494.1	WP_011617242.1
WP_081050203.1	WP_037025449.1	WP_011617753.1	WP_011617493.1	WP_011617241.1
WP_081050152.1	WP_037025439.1	WP_011617752.1	WP_011617492.1	WP_011617240.1
WP_081050151.1	WP_037025429.1	WP_011617750.1	WP_011617490.1	WP_011617239.1
WP_063834067.1	WP_037025021.1	WP_011617749.1	WP_011617489.1	WP_011617238.1

Appendix B

WP_051398616.1	WP_037025012.1	WP_011617747.1	WP_011617487.1	WP_011617237.1
WP_051398615.1	WP_037024995.1	WP_011617746.1	WP_011617486.1	WP_011617236.1
WP_051398612.1	WP_037024847.1	WP_011617745.1	WP_011617485.1	WP_011617235.1
WP_051398611.1	WP_037024831.1	WP_011617744.1	WP_011617484.1	WP_011617232.1
WP_051398610.1	WP_037024810.1	WP_011617743.1	WP_011617483.1	WP_011617230.1
WP_051398609.1	WP_037024770.1	WP_011617740.1	WP_011617482.1	WP_011617228.1
WP_051398608.1	WP_037024735.1	WP_011617739.1	WP_011617480.1	WP_011617226.1
WP_051398604.1	WP_037024693.1	WP_011617737.1	WP_011617479.1	WP_011617225.1
WP_051398602.1	WP_037024323.1	WP_011617736.1	WP_011617478.1	WP_011617224.1
WP_051398601.1	WP_037024254.1	WP_011617734.1	WP_011617474.1	WP_011617223.1
WP_051398600.1	WP_037024235.1	WP_011617733.1	WP_011617473.1	WP_011617222.1
WP_051398593.1	WP_037024190.1	WP_011617732.1	WP_011617471.1	WP_011617221.1
WP_051398592.1	WP_037024181.1	WP_011617731.1	WP_011617470.1	WP_011617220.1
WP_051398589.1	WP_037024046.1	WP_011617729.1	WP_011617469.1	WP_011617219.1
WP_051398587.1	WP_037024037.1	WP_011617728.1	WP_011617468.1	WP_011617218.1
WP_051398586.1	WP_037024008.1	WP_011617726.1	WP_011617467.1	WP_011617217.1
WP_051398576.1	WP_037023720.1	WP_011617724.1	WP_011617466.1	WP_011617216.1
WP_051398574.1	WP_037023717.1	WP_011617723.1	WP_011617464.1	WP_011617215.1
WP_051398573.1	WP_037023464.1	WP_011617722.1	WP_011617463.1	WP_011617214.1
WP_051398566.1	WP_037023447.1	WP_011617721.1	WP_011617462.1	WP_011617213.1
WP_051398564.1	WP_037023437.1	WP_011617720.1	WP_011617461.1	WP_011617212.1
WP_051398563.1	WP_037023372.1	WP_011617719.1	WP_011617459.1	WP_011617211.1
WP_051398557.1	WP_037023370.1	WP_011617718.1	WP_011617458.1	WP_011617209.1
WP_051398556.1	WP_037023361.1	WP_011617717.1	WP_011617457.1	WP_011617208.1
WP_051398555.1	WP_037023349.1	WP_011617715.1	WP_011617455.1	WP_011617207.1
WP_051398554.1	WP_035823937.1	WP_011617714.1	WP_011617454.1	WP_011617206.1
WP_051398553.1	WP_035823310.1	WP_011617713.1	WP_011617453.1	WP_011617205.1
WP_051398552.1	WP_026200869.1	WP_011617710.1	WP_011617452.1	WP_011617204.1
WP_051398547.1	WP_018315435.1	WP_011617708.1	WP_011617451.1	WP_011617203.1
WP_041688591.1	WP_013953596.1	WP_011617707.1	WP_011617450.1	WP_011617202.1
WP_041688575.1	WP_011617972.1	WP_011617706.1	WP_011617449.1	WP_011617200.1
WP_041688572.1	WP_011617970.1	WP_011617705.1	WP_011617448.1	WP_011617199.1
WP_041688562.1	WP_011617969.1	WP_011617704.1	WP_011617447.1	WP_011617198.1
WP_041688560.1	WP_011617967.1	WP_011617703.1	WP_011617446.1	WP_011617197.1
WP_041688558.1	WP_011617966.1	WP_011617702.1	WP_011617444.1	WP_011617196.1
WP_041688550.1	WP_011617964.1	WP_011617701.1	WP_011617443.1	WP_011617195.1
WP_041688546.1	WP_011617963.1	WP_011617700.1	WP_011617440.1	WP_011617193.1
WP_041688544.1	WP_011617962.1	WP_011617699.1	WP_011617438.1	WP_011617192.1
WP_041688521.1	WP_011617961.1	WP_011617698.1	WP_011617437.1	WP_011617191.1
WP_041688513.1	WP_011617960.1	WP_011617697.1	WP_011617436.1	WP_011617190.1
WP_041688507.1	WP_011617959.1	WP_011617694.1	WP_011617435.1	WP_011617189.1
WP_041688505.1	WP_011617958.1	WP_011617693.1	WP_011617433.1	WP_011617188.1

Appendix B

WP_041688501.1	WP_011617957.1	WP_011617692.1	WP_011617432.1	WP_011617187.1
WP_041688485.1	WP_011617955.1	WP_011617691.1	WP_011617431.1	WP_011617185.1
WP_041688483.1	WP_011617954.1	WP_011617688.1	WP_011617429.1	WP_011617184.1
WP_041688461.1	WP_011617953.1	WP_011617687.1	WP_011617428.1	WP_011617183.1
WP_041688449.1	WP_011617952.1	WP_011617686.1	WP_011617427.1	WP_011617182.1
WP_041688444.1	WP_011617950.1	WP_011617685.1	WP_011617426.1	WP_011617181.1

WP_011617180.1	WP_011616916.1	WP_011616679.1	WP_011616419.1	WP_010814120.1
WP_011617179.1	WP_011616915.1	WP_011616678.1	WP_011616418.1	WP_010814118.1
WP_011617178.1	WP_011616914.1	WP_011616677.1	WP_011616417.1	WP_010814098.1
WP_011617176.1	WP_011616913.1	WP_011616676.1	WP_011616416.1	WP_010814096.1
WP_011617175.1	WP_011616912.1	WP_011616675.1	WP_011616415.1	WP_010814090.1
WP_011617172.1	WP_011616911.1	WP_011616673.1	WP_011616414.1	WP_010814088.1
WP_011617171.1	WP_011616909.1	WP_011616672.1	WP_011616413.1	WP_010814084.1
WP_011617168.1	WP_011616908.1	WP_011616670.1	WP_011616412.1	WP_010814082.1
WP_011617167.1	WP_011616907.1	WP_011616669.1	WP_011616411.1	WP_010814081.1
WP_011617166.1	WP_011616906.1	WP_011616668.1	WP_011616410.1	WP_010814080.1
WP_011617164.1	WP_011616905.1	WP_011616667.1	WP_011616409.1	WP_010814077.1
WP_011617163.1	WP_011616904.1	WP_011616665.1	WP_011616407.1	WP_010814070.1
WP_011617161.1	WP_011616903.1	WP_011616664.1	WP_011616405.1	WP_010814069.1
WP_011617160.1	WP_011616901.1	WP_011616661.1	WP_011616404.1	WP_010814067.1
WP_011617158.1	WP_011616900.1	WP_011616660.1	WP_011616403.1	WP_010814065.1
WP_011617157.1	WP_011616899.1	WP_011616659.1	WP_011616402.1	WP_010814063.1
WP_011617156.1	WP_011616898.1	WP_011616658.1	WP_011616401.1	WP_010814062.1
WP_011617155.1	WP_011616896.1	WP_011616657.1	WP_011616400.1	WP_010814061.1
WP_011617154.1	WP_011616895.1	WP_011616656.1	WP_011616399.1	WP_010814059.1
WP_011617153.1	WP_011616894.1	WP_011616655.1	WP_011616398.1	WP_010814058.1
WP_011617152.1	WP_011616893.1	WP_011616654.1	WP_011616396.1	WP_010814057.1
WP_011617149.1	WP_011616892.1	WP_011616653.1	WP_011616395.1	WP_010814056.1
WP_011617148.1	WP_011616891.1	WP_011616651.1	WP_011616394.1	WP_010814055.1
WP_011617146.1	WP_011616890.1	WP_011616650.1	WP_011616392.1	WP_010814054.1
WP_011617145.1	WP_011616889.1	WP_011616649.1	WP_011616391.1	WP_010814051.1
WP_011617144.1	WP_011616888.1	WP_011616648.1	WP_011616390.1	WP_010814050.1
WP_011617142.1	WP_011616887.1	WP_011616647.1	WP_011616389.1	WP_010814049.1
WP_011617141.1	WP_011616886.1	WP_011616645.1	WP_011616388.1	WP_010814048.1
WP_011617140.1	WP_011616885.1	WP_011616644.1	WP_011616387.1	WP_010814047.1
WP_011617139.1	WP_011616884.1	WP_011616642.1	WP_011616385.1	WP_010814038.1
WP_011617137.1	WP_011616883.1	WP_011616641.1	WP_011616384.1	WP_010814037.1
WP_011617135.1	WP_011616882.1	WP_011616640.1	WP_011616383.1	WP_010814035.1
WP_011617134.1	WP_011616881.1	WP_011616639.1	WP_011616382.1	WP_010814032.1
WP_011617133.1	WP_011616880.1	WP_011616638.1	WP_011616381.1	WP_010814028.1

Appendix B

WP_011617132.1	WP_011616879.1	WP_011616637.1	WP_011616380.1	WP_010814027.1
WP_011617131.1	WP_011616876.1	WP_011616636.1	WP_011616379.1	WP_010814024.1
WP_011617130.1	WP_011616875.1	WP_011616635.1	WP_011616378.1	WP_010814013.1
WP_011617129.1	WP_011616873.1	WP_011616634.1	WP_011616377.1	WP_010814009.1
WP_011617128.1	WP_011616872.1	WP_011616633.1	WP_011616374.1	WP_010814008.1
WP_011617127.1	WP_011616871.1	WP_011616632.1	WP_011616372.1	WP_010814001.1
WP_011617126.1	WP_011616870.1	WP_011616631.1	WP_011616371.1	WP_010814000.1
WP_011617125.1	WP_011616869.1	WP_011616630.1	WP_011616370.1	WP_010813996.1
WP_011617124.1	WP_011616868.1	WP_011616629.1	WP_011616368.1	WP_010813994.1
WP_011617123.1	WP_011616866.1	WP_011616628.1	WP_011616367.1	WP_010813991.1
WP_011617122.1	WP_011616865.1	WP_011616627.1	WP_011616366.1	WP_010813984.1
WP_011617121.1	WP_011616864.1	WP_011616626.1	WP_011616365.1	WP_010813979.1
WP_011617119.1	WP_011616863.1	WP_011616625.1	WP_011616364.1	WP_010813977.1
WP_011617118.1	WP_011616862.1	WP_011616624.1	WP_011616363.1	WP_010813973.1
WP_011617117.1	WP_011616861.1	WP_011616622.1	WP_011616362.1	WP_010813971.1
WP_011617116.1	WP_011616860.1	WP_011616621.1	WP_011616361.1	WP_010813969.1
WP_011617114.1	WP_011616859.1	WP_011616620.1	WP_011616360.1	WP_010813961.1
WP_011617113.1	WP_011616858.1	WP_011616619.1	WP_011616359.1	WP_010813959.1
WP_011617111.1	WP_011616857.1	WP_011616618.1	WP_011616358.1	WP_010813951.1
WP_011617110.1	WP_011616856.1	WP_011616617.1	WP_011616357.1	WP_010813950.1
WP_011617109.1	WP_011616855.1	WP_011616616.1	WP_011616356.1	WP_010813949.1
WP_011617108.1	WP_011616854.1	WP_011616615.1	WP_011616355.1	WP_010813948.1
WP_011617106.1	WP_011616852.1	WP_011616614.1	WP_011616354.1	WP_010813947.1
WP_011617105.1	WP_011616851.1	WP_011616613.1	WP_011616353.1	WP_010813946.1
WP_011617104.1	WP_011616849.1	WP_011616612.1	WP_011616352.1	WP_010813533.1
WP_011617103.1	WP_011616847.1	WP_011616611.1	WP_011616351.1	WP_010813531.1
WP_011617102.1	WP_011616846.1	WP_011616610.1	WP_011616349.1	WP_010813530.1
WP_011617100.1	WP_011616843.1	WP_011616608.1	WP_011616348.1	WP_010813529.1
WP_011617099.1	WP_011616841.1	WP_011616607.1	WP_011616347.1	WP_010813525.1
WP_011617098.1	WP_011616840.1	WP_011616606.1	WP_011616346.1	WP_010813519.1
WP_011617097.1	WP_011616839.1	WP_011616605.1	WP_011616345.1	WP_010813517.1
WP_011617096.1	WP_011616837.1	WP_011616604.1	WP_011616342.1	WP_010813516.1
WP_011617095.1	WP_011616836.1	WP_011616602.1	WP_011616341.1	WP_010813515.1
WP_011617091.1	WP_011616835.1	WP_011616598.1	WP_011616340.1	WP_010813513.1
WP_011617090.1	WP_011616834.1	WP_011616588.1	WP_011616339.1	WP_010813511.1
WP_011617089.1	WP_011616832.1	WP_011616587.1	WP_011616338.1	WP_010813507.1
WP_011617085.1	WP_011616831.1	WP_011616585.1	WP_011616337.1	WP_010813506.1
WP_011617084.1	WP_011616830.1	WP_011616581.1	WP_011616336.1	WP_010813503.1
WP_011617082.1	WP_011616829.1	WP_011616580.1	WP_011616335.1	WP_010813499.1
WP_011617081.1	WP_011616828.1	WP_011616579.1	WP_011616334.1	WP_010813494.1
WP_011617080.1	WP_011616827.1	WP_011616578.1	WP_011616333.1	WP_010813492.1
WP_011617079.1	WP_011616826.1	WP_011616577.1	WP_011616332.1	WP_010813487.1

Appendix B

WP_011617078.1	WP_011616825.1	WP_011616575.1	WP_011616331.1	WP_010813486.1
WP_011617077.1	WP_011616824.1	WP_011616574.1	WP_011616330.1	WP_010813481.1
WP_011617075.1	WP_011616823.1	WP_011616572.1	WP_011616329.1	WP_010813479.1
WP_011617074.1	WP_011616822.1	WP_011616571.1	WP_011616328.1	WP_010813419.1
WP_011617073.1	WP_011616821.1	WP_011616570.1	WP_011616327.1	WP_010813418.1
WP_011617072.1	WP_011616818.1	WP_011616568.1	WP_011616326.1	WP_010813417.1
WP_011617071.1	WP_011616817.1	WP_011616567.1	WP_011616325.1	WP_010813416.1
WP_011617070.1	WP_011616816.1	WP_011616566.1	WP_011616322.1	WP_010813415.1
WP_011617069.1	WP_011616814.1	WP_011616565.1	WP_011616321.1	WP_010813414.1
WP_011617068.1	WP_011616813.1	WP_011616564.1	WP_011616320.1	WP_010813413.1
WP_011617067.1	WP_011616812.1	WP_011616563.1	WP_011616319.1	WP_010813403.1
WP_011617066.1	WP_011616811.1	WP_011616562.1	WP_011616318.1	WP_010813402.1
WP_011617064.1	WP_011616810.1	WP_011616560.1	WP_011616317.1	WP_010813395.1
WP_011617062.1	WP_011616809.1	WP_011616559.1	WP_011616315.1	WP_010813392.1
WP_011617060.1	WP_011616808.1	WP_011616558.1	WP_011616314.1	WP_010813390.1
WP_011617059.1	WP_011616807.1	WP_011616557.1	WP_011616313.1	WP_010813389.1
WP_011617058.1	WP_011616806.1	WP_011616556.1	WP_011616312.1	WP_010813388.1
WP_011617057.1	WP_011616804.1	WP_011616555.1	WP_011616309.1	WP_010813386.1
WP_011617056.1	WP_011616803.1	WP_011616553.1	WP_011616307.1	WP_010813380.1
WP_011617055.1	WP_011616802.1	WP_011616552.1	WP_011616306.1	WP_010813374.1
WP_011617054.1	WP_011616801.1	WP_011616550.1	WP_011616305.1	WP_010813371.1
WP_011617052.1	WP_011616800.1	WP_011616549.1	WP_011616304.1	WP_010813368.1
WP_011617051.1	WP_011616799.1	WP_011616548.1	WP_011616303.1	WP_010813250.1
WP_011617049.1	WP_011616798.1	WP_011616546.1	WP_011616302.1	WP_010813243.1
WP_011617048.1	WP_011616797.1	WP_011616545.1	WP_011616299.1	WP_010813242.1
WP_011617047.1	WP_011616796.1	WP_011616544.1	WP_011616296.1	WP_010813237.1
WP_011617046.1	WP_011616795.1	WP_011616543.1	WP_011615250.1	WP_010812869.1
WP_011617045.1	WP_011616794.1	WP_011616542.1	WP_010814599.1	WP_010812868.1
WP_011617044.1	WP_011616792.1	WP_011616541.1	WP_010814596.1	WP_010812866.1
WP_011617043.1	WP_011616791.1	WP_011616540.1	WP_010814590.1	WP_010812859.1
WP_011617040.1	WP_011616790.1	WP_011616539.1	WP_010814588.1	WP_010812856.1
WP_011617039.1	WP_011616789.1	WP_011616538.1	WP_010814584.1	WP_010812855.1
WP_011617038.1	WP_011616787.1	WP_011616536.1	WP_010814581.1	WP_010812853.1
WP_011617037.1	WP_011616786.1	WP_011616533.1	WP_010814580.1	WP_010812852.1
WP_011617036.1	WP_011616785.1	WP_011616532.1	WP_010814579.1	WP_010812851.1
WP_011617035.1	WP_011616784.1	WP_011616531.1	WP_010814578.1	WP_010812848.1
WP_011617034.1	WP_011616783.1	WP_011616528.1	WP_010814571.1	WP_010812847.1
WP_011617032.1	WP_011616782.1	WP_011616527.1	WP_010814570.1	WP_010812846.1
WP_011617031.1	WP_011616780.1	WP_011616526.1	WP_010814566.1	WP_010812844.1
WP_011617030.1	WP_011616779.1	WP_011616525.1	WP_010814565.1	WP_010812842.1
WP_011617029.1	WP_011616778.1	WP_011616522.1	WP_010814564.1	WP_010812839.1
WP_011617028.1	WP_011616776.1	WP_011616521.1	WP_010814563.1	WP_010812838.1

Appendix B

WP_011617027.1	WP_011616775.1	WP_011616520.1	WP_010814562.1	WP_010812837.1
WP_011617025.1	WP_011616774.1	WP_011616519.1	WP_010814561.1	WP_010812834.1
WP_011617024.1	WP_011616773.1	WP_011616518.1	WP_010814559.1	WP_010812833.1
WP_011617023.1	WP_011616772.1	WP_011616517.1	WP_010814558.1	WP_010812832.1
WP_011617021.1	WP_011616770.1	WP_011616516.1	WP_010814556.1	WP_010812831.1
WP_011617020.1	WP_011616768.1	WP_011616514.1	WP_010814554.1	WP_010812830.1
WP_011617018.1	WP_011616767.1	WP_011616513.1	WP_010814553.1	WP_010812829.1
WP_011617017.1	WP_011616765.1	WP_011616512.1	WP_010814552.1	WP_010812825.1
WP_011617016.1	WP_011616764.1	WP_011616511.1	WP_010814551.1	WP_010812824.1
WP_011617015.1	WP_011616763.1	WP_011616510.1	WP_010814545.1	WP_010812822.1
WP_011617014.1	WP_011616762.1	WP_011616509.1	WP_010814544.1	WP_010812821.1
WP_011617013.1	WP_011616761.1	WP_011616508.1	WP_010814542.1	WP_010812820.1
WP_011617012.1	WP_011616760.1	WP_011616507.1	WP_010814538.1	WP_010812817.1
WP_011617011.1	WP_011616759.1	WP_011616506.1	WP_010814536.1	WP_010812815.1
WP_011617010.1	WP_011616758.1	WP_011616505.1	WP_010814535.1	WP_010812814.1
WP_011617009.1	WP_011616757.1	WP_011616504.1	WP_010814534.1	WP_010812813.1
WP_011617008.1	WP_011616756.1	WP_011616503.1	WP_010814533.1	WP_010812812.1
WP_011617005.1	WP_011616755.1	WP_011616502.1	WP_010814531.1	WP_010812811.1
WP_011617003.1	WP_011616754.1	WP_011616501.1	WP_010814528.1	WP_010812809.1
WP_011617002.1	WP_011616753.1	WP_011616500.1	WP_010814521.1	WP_010812808.1
WP_011617001.1	WP_011616752.1	WP_011616499.1	WP_010814520.1	WP_010812807.1
WP_011616996.1	WP_011616751.1	WP_011616498.1	WP_010814518.1	WP_010812799.1
WP_011616995.1	WP_011616750.1	WP_011616497.1	WP_010814517.1	WP_010812798.1
WP_011616994.1	WP_011616749.1	WP_011616496.1	WP_010814516.1	WP_010812796.1
WP_011616993.1	WP_011616748.1	WP_011616494.1	WP_010814506.1	WP_010812795.1
WP_011616992.1	WP_011616747.1	WP_011616493.1	WP_010814505.1	WP_010812792.1
WP_011616991.1	WP_011616746.1	WP_011616492.1	WP_010814503.1	WP_010812789.1
WP_011616990.1	WP_011616745.1	WP_011616491.1	WP_010814502.1	WP_010812788.1
WP_011616989.1	WP_011616743.1	WP_011616490.1	WP_010814500.1	WP_010812787.1
WP_011616985.1	WP_011616742.1	WP_011616489.1	WP_010814498.1	WP_010812783.1
WP_011616983.1	WP_011616741.1	WP_011616488.1	WP_010814497.1	WP_010812782.1
WP_011616982.1	WP_011616740.1	WP_011616487.1	WP_010814494.1	WP_010812781.1
WP_011616980.1	WP_011616739.1	WP_011616486.1	WP_010814493.1	WP_010812775.1
WP_011616979.1	WP_011616738.1	WP_011616485.1	WP_010814491.1	WP_010812772.1
WP_011616977.1	WP_011616737.1	WP_011616484.1	WP_010814488.1	WP_010812769.1
WP_011616973.1	WP_011616736.1	WP_011616480.1	WP_010814487.1	WP_010812767.1
WP_011616972.1	WP_011616735.1	WP_011616479.1	WP_010814484.1	WP_010812764.1
WP_011616970.1	WP_011616734.1	WP_011616478.1	WP_010814477.1	WP_010812760.1
WP_011616968.1	WP_011616733.1	WP_011616477.1	WP_010814475.1	WP_010812758.1
WP_011616967.1	WP_011616732.1	WP_011616476.1	WP_010814473.1	WP_010812756.1
WP_011616966.1	WP_011616731.1	WP_011616473.1	WP_010814472.1	WP_010812754.1
WP_011616965.1	WP_011616730.1	WP_011616472.1	WP_010814471.1	WP_010812752.1

Appendix B

WP_011616964.1	WP_011616729.1	WP_011616468.1	WP_010814470.1	WP_010812751.1
WP_011616963.1	WP_011616728.1	WP_011616467.1	WP_010814469.1	WP_010812746.1
WP_011616962.1	WP_011616727.1	WP_011616466.1	WP_010814465.1	WP_010812745.1
WP_011616961.1	WP_011616726.1	WP_011616465.1	WP_010814459.1	WP_010812743.1
WP_011616960.1	WP_011616724.1	WP_011616464.1	WP_010814457.1	WP_010812741.1
WP_011616957.1	WP_011616722.1	WP_011616463.1	WP_010814456.1	WP_010812739.1
WP_011616956.1	WP_011616721.1	WP_011616462.1	WP_010814455.1	WP_010812738.1
WP_011616955.1	WP_011616720.1	WP_011616459.1	WP_010814452.1	WP_010812735.1
WP_011616954.1	WP_011616719.1	WP_011616458.1	WP_010814449.1	WP_010812734.1
WP_011616953.1	WP_011616718.1	WP_011616457.1	WP_010814242.1	WP_010812733.1
WP_011616952.1	WP_011616717.1	WP_011616456.1	WP_010814240.1	WP_010812729.1
WP_011616951.1	WP_011616715.1	WP_011616455.1	WP_010814239.1	WP_010812722.1
WP_011616948.1	WP_011616714.1	WP_011616453.1	WP_010814237.1	WP_010812719.1
WP_011616947.1	WP_011616713.1	WP_011616451.1	WP_010814236.1	WP_010812713.1
WP_011616945.1	WP_011616712.1	WP_011616450.1	WP_010814234.1	WP_010812711.1
WP_011616944.1	WP_011616711.1	WP_011616449.1	WP_010814233.1	WP_010812710.1
WP_011616943.1	WP_011616710.1	WP_011616448.1	WP_010814229.1	WP_010812708.1
WP_011616942.1	WP_011616709.1	WP_011616446.1	WP_010814227.1	WP_010812707.1
WP_011616941.1	WP_011616708.1	WP_011616445.1	WP_010814226.1	WP_010812706.1
WP_011616940.1	WP_011616706.1	WP_011616444.1	WP_010814168.1	WP_010812705.1
WP_011616939.1	WP_011616705.1	WP_011616443.1	WP_010814167.1	WP_010812673.1
WP_011616938.1	WP_011616704.1	WP_011616442.1	WP_010814166.1	WP_010812672.1
WP_011616937.1	WP_011616703.1	WP_011616441.1	WP_010814165.1	WP_010812671.1
WP_011616936.1	WP_011616702.1	WP_011616440.1	WP_010814164.1	WP_010812670.1
WP_011616935.1	WP_011616701.1	WP_011616439.1	WP_010814163.1	WP_010812668.1
WP_011616933.1	WP_011616700.1	WP_011616438.1	WP_010814159.1	WP_010812667.1
WP_011616932.1	WP_011616699.1	WP_011616437.1	WP_010814158.1	WP_010812666.1
WP_011616930.1	WP_011616698.1	WP_011616434.1	WP_010814144.1	WP_010812665.1
WP_011616929.1	WP_011616697.1	WP_011616433.1	WP_010814143.1	WP_010812663.1
WP_011616928.1	WP_011616696.1	WP_011616432.1	WP_010814141.1	WP_010812662.1
WP_011616927.1	WP_011616694.1	WP_011616430.1	WP_010814139.1	WP_010812652.1
WP_011616926.1	WP_011616693.1	WP_011616429.1	WP_010814138.1	WP_010812637.1
WP_011616925.1	WP_011616692.1	WP_011616428.1	WP_010814136.1	WP_010812634.1
WP_011616924.1	WP_011616691.1	WP_011616427.1	WP_010814127.1	WP_010812630.1
WP_011616922.1	WP_011616690.1	WP_011616425.1	WP_010814126.1	WP_010812626.1
WP_011616921.1	WP_011616689.1	WP_011616424.1	WP_010814125.1	WP_010812625.1
WP_011616920.1	WP_011616688.1	WP_011616423.1	WP_010814124.1	WP_010812613.1
WP_011616919.1	WP_011616687.1	WP_011616422.1	WP_010814123.1	WP_010812612.1
WP_011616918.1	WP_011616684.1	WP_011616421.1	WP_010814122.1	WP_010812610.1
WP_011616917.1	WP_011616680.1	WP_011616420.1	WP_010814121.1	WP_010812609.1

Appendix B

WP_010812603.1	WP_010810991.1	WP_010810283.1	WP_010809265.1
WP_010812599.1	WP_010810985.1	WP_010810281.1	WP_010809263.1
WP_010812595.1	WP_010810981.1	WP_010810279.1	WP_010809262.1
WP_010812593.1	WP_010810979.1	WP_010810278.1	WP_010809259.1
WP_010812588.1	WP_010810973.1	WP_010810276.1	WP_010809257.1
WP_010812586.1	WP_010810971.1	WP_010810275.1	WP_010809256.1
WP_010812585.1	WP_010810967.1	WP_010810271.1	WP_010809253.1
WP_010812584.1	WP_010810966.1	WP_010810269.1	WP_010809252.1
WP_010812576.1	WP_010810964.1	WP_010810268.1	WP_008646230.1
WP_010812572.1	WP_010810959.1	WP_010810266.1	WP_006159056.1
WP_010812571.1	WP_010810958.1	WP_010810265.1	
WP_010812567.1	WP_010810955.1	WP_010810264.1	
WP_010812563.1	WP_010810954.1	WP_010810263.1	
WP_010812559.1	WP_010810953.1	WP_010810260.1	
WP_010812556.1	WP_010810952.1	WP_010810257.1	
WP_010812548.1	WP_010810950.1	WP_010810256.1	
WP_010812546.1	WP_010810945.1	WP_010810251.1	
WP_010812541.1	WP_010810938.1	WP_010810250.1	
WP_010812539.1	WP_010810936.1	WP_010810249.1	
WP_010812533.1	WP_010810935.1	WP_010810246.1	
WP_010812532.1	WP_010810931.1	WP_010810245.1	
WP_010812531.1	WP_010810926.1	WP_010810244.1	
WP_010812519.1	WP_010810921.1	WP_010810242.1	
WP_010812518.1	WP_010810918.1	WP_010810239.1	
WP_010812516.1	WP_010810917.1	WP_010810237.1	
WP_010812515.1	WP_010810916.1	WP_010810236.1	
WP_010812514.1	WP_010810912.1	WP_010810235.1	
WP_010812511.1	WP_010810907.1	WP_010810230.1	
WP_010812510.1	WP_010810906.1	WP_010810229.1	
WP_010812509.1	WP_010810904.1	WP_010810228.1	
WP_010812504.1	WP_010810903.1	WP_010810226.1	
WP_010812502.1	WP_010810902.1	WP_010810224.1	
WP_010812500.1	WP_010810900.1	WP_010810218.1	
WP_010812498.1	WP_010810897.1	WP_010810216.1	
WP_010812497.1	WP_010810894.1	WP_010810215.1	
WP_010812495.1	WP_010810893.1	WP_010810212.1	
WP_010812494.1	WP_010810892.1	WP_010810209.1	
WP_010812493.1	WP_010810891.1	WP_010810208.1	
WP_010812490.1	WP_010810889.1	WP_010810206.1	
WP_010812488.1	WP_010810887.1	WP_010810205.1	
WP_010812486.1	WP_010810883.1	WP_010810202.1	
WP_010812485.1	WP_010810881.1	WP_010810201.1	

Appendix B

WP_010812471.1	WP_010810880.1	WP_010810198.1	
WP_010812468.1	WP_010810879.1	WP_010810192.1	
WP_010812465.1	WP_010810878.1	WP_010810181.1	
WP_010812461.1	WP_010810873.1	WP_010810173.1	
WP_010812460.1	WP_010810870.1	WP_010810172.1	
WP_010812454.1	WP_010810867.1	WP_010810161.1	
WP_010812450.1	WP_010810866.1	WP_010810153.1	
WP_010812447.1	WP_010810863.1	WP_010810151.1	
WP_010812445.1	WP_010810861.1	WP_010810150.1	
WP_010812442.1	WP_010810860.1	WP_010810148.1	
WP_010812441.1	WP_010810858.1	WP_010809863.1	
WP_010812439.1	WP_010810857.1	WP_010809862.1	
WP_010812436.1	WP_010810853.1	WP_010809861.1	
WP_010812435.1	WP_010810842.1	WP_010809859.1	
WP_010812434.1	WP_010810836.1	WP_010809858.1	
WP_010812432.1	WP_010810834.1	WP_010809856.1	
WP_010812430.1	WP_010810832.1	WP_010809854.1	
WP_010812428.1	WP_010810830.1	WP_010809853.1	
WP_010812427.1	WP_010810826.1	WP_010809850.1	
WP_010812426.1	WP_010810825.1	WP_010809845.1	
WP_010812425.1	WP_010810823.1	WP_010809843.1	
WP_010812421.1	WP_010810820.1	WP_010809841.1	
WP_010812420.1	WP_010810819.1	WP_010809834.1	
WP_010812418.1	WP_010810818.1	WP_010809833.1	
WP_010812413.1	WP_010810813.1	WP_010809831.1	
WP_010812410.1	WP_010810810.1	WP_010809830.1	
WP_010812408.1	WP_010810809.1	WP_010809826.1	
WP_010812407.1	WP_010810807.1	WP_010809819.1	
WP_010812405.1	WP_010810784.1	WP_010809818.1	
WP_010811720.1	WP_010810779.1	WP_010809815.1	
WP_010811718.1	WP_010810774.1	WP_010809813.1	
WP_010811717.1	WP_010810773.1	WP_010809809.1	
WP_010811716.1	WP_010810766.1	WP_010809807.1	
WP_010811713.1	WP_010810761.1	WP_010809806.1	
WP_010811712.1	WP_010810757.1	WP_010809805.1	
WP_010811709.1	WP_010810755.1	WP_010809803.1	
WP_010811708.1	WP_010810754.1	WP_010809799.1	
WP_010811703.1	WP_010810752.1	WP_010809797.1	
WP_010811700.1	WP_010810750.1	WP_010809792.1	
WP_010811697.1	WP_010810746.1	WP_010809789.1	
WP_010811693.1	WP_010810745.1	WP_010809788.1	
WP_010811692.1	WP_010810744.1	WP_010809786.1	

Appendix B

WP_010811691.1	WP_010810742.1	WP_010809785.1	
WP_010811690.1	WP_010810735.1	WP_010809784.1	
WP_010811687.1	WP_010810734.1	WP_010809783.1	
WP_010811686.1	WP_010810731.1	WP_010809779.1	
WP_010811682.1	WP_010810723.1	WP_010809778.1	
WP_010811681.1	WP_010810722.1	WP_010809775.1	
WP_010811679.1	WP_010810720.1	WP_010809774.1	
WP_010811677.1	WP_010810719.1	WP_010809773.1	
WP_010811676.1	WP_010810714.1	WP_010809772.1	
WP_010811674.1	WP_010810713.1	WP_010809769.1	
WP_010811673.1	WP_010810707.1	WP_010809765.1	
WP_010811672.1	WP_010810705.1	WP_010809762.1	
WP_010811671.1	WP_010810704.1	WP_010809760.1	
WP_010811670.1	WP_010810702.1	WP_010809756.1	
WP_010811669.1	WP_010810699.1	WP_010809754.1	
WP_010811666.1	WP_010810697.1	WP_010809753.1	
WP_010811662.1	WP_010810694.1	WP_010809749.1	
WP_010811661.1	WP_010810690.1	WP_010809741.1	
WP_010811660.1	WP_010810688.1	WP_010809738.1	
WP_010811658.1	WP_010810687.1	WP_010809737.1	
WP_010811657.1	WP_010810682.1	WP_010809735.1	
WP_010811654.1	WP_010810681.1	WP_010809734.1	
WP_010811652.1	WP_010810675.1	WP_010809729.1	
WP_010811651.1	WP_010810674.1	WP_010809728.1	
WP_010811649.1	WP_010810671.1	WP_010809727.1	
WP_010811648.1	WP_010810670.1	WP_010809726.1	
WP_010811645.1	WP_010810669.1	WP_010809725.1	
WP_010811641.1	WP_010810666.1	WP_010809724.1	
WP_010811640.1	WP_010810665.1	WP_010809720.1	
WP_010811638.1	WP_010810659.1	WP_010809719.1	
WP_010811637.1	WP_010810657.1	WP_010809714.1	
WP_010811636.1	WP_010810656.1	WP_010809704.1	
WP_010811635.1	WP_010810654.1	WP_010809703.1	
WP_010811633.1	WP_010810653.1	WP_010809702.1	
WP_010811632.1	WP_010810649.1	WP_010809701.1	
WP_010811630.1	WP_010810644.1	WP_010809687.1	
WP_010811626.1	WP_010810642.1	WP_010809684.1	
WP_010811624.1	WP_010810455.1	WP_010809683.1	
WP_010811200.1	WP_010810454.1	WP_010809673.1	
WP_010811197.1	WP_010810452.1	WP_010809672.1	
WP_010811195.1	WP_010810451.1	WP_010809665.1	
WP_010811194.1	WP_010810450.1	WP_010809663.1	

Appendix B

WP_010811190.1	WP_010810448.1	WP_010809661.1	
WP_010811188.1	WP_010810446.1	WP_010809660.1	
WP_010811186.1	WP_010810445.1	WP_010809659.1	
WP_010811185.1	WP_010810436.1	WP_010809656.1	
WP_010811182.1	WP_010810425.1	WP_010809655.1	
WP_010811180.1	WP_010810424.1	WP_010809653.1	
WP_010811178.1	WP_010810421.1	WP_010809650.1	
WP_010811176.1	WP_010810420.1	WP_010809648.1	
WP_010811174.1	WP_010810419.1	WP_010809646.1	
WP_010811172.1	WP_010810418.1	WP_010809643.1	
WP_010811168.1	WP_010810417.1	WP_010809642.1	
WP_010811167.1	WP_010810415.1	WP_010809639.1	
WP_010811156.1	WP_010810411.1	WP_010809636.1	
WP_010811152.1	WP_010810410.1	WP_010809635.1	
WP_010811149.1	WP_010810409.1	WP_010809634.1	
WP_010811147.1	WP_010810406.1	WP_010809624.1	
WP_010811146.1	WP_010810403.1	WP_010809622.1	
WP_010811144.1	WP_010810401.1	WP_010809618.1	
WP_010811139.1	WP_010810400.1	WP_010809615.1	
WP_010811137.1	WP_010810398.1	WP_010809614.1	
WP_010811135.1	WP_010810396.1	WP_010809411.1	
WP_010811134.1	WP_010810395.1	WP_010809408.1	
WP_010811132.1	WP_010810394.1	WP_010809407.1	
WP_010811131.1	WP_010810392.1	WP_010809405.1	
WP_010811128.1	WP_010810389.1	WP_010809402.1	
WP_010811116.1	WP_010810388.1	WP_010809400.1	
WP_010811115.1	WP_010810387.1	WP_010809399.1	
WP_010811109.1	WP_010810386.1	WP_010809397.1	
WP_010811108.1	WP_010810385.1	WP_010809389.1	
WP_010811105.1	WP_010810381.1	WP_010809388.1	
WP_010811104.1	WP_010810380.1	WP_010809386.1	
WP_010811101.1	WP_010810377.1	WP_010809381.1	
WP_010811100.1	WP_010810376.1	WP_010809376.1	
WP_010811095.1	WP_010810375.1	WP_010809373.1	
WP_010811091.1	WP_010810373.1	WP_010809372.1	
WP_010811090.1	WP_010810371.1	WP_010809371.1	
WP_010811086.1	WP_010810370.1	WP_010809369.1	
WP_010811084.1	WP_010810369.1	WP_010809367.1	
WP_010811083.1	WP_010810366.1	WP_010809364.1	
WP_010811082.1	WP_010810364.1	WP_010809362.1	
WP_010811081.1	WP_010810363.1	WP_010809361.1	
WP_010811077.1	WP_010810360.1	WP_010809360.1	

WP_010811074.1	WP_010810358.1	WP_010809358.1	
WP_010811071.1	WP_010810356.1	WP_010809356.1	
WP_010811070.1	WP_010810355.1	WP_010809352.1	
WP_010811069.1	WP_010810353.1	WP_010809350.1	
WP_010811068.1	WP_010810347.1	WP_010809344.1	
WP_010811067.1	WP_010810343.1	WP_010809335.1	
WP_010811062.1	WP_010810337.1	WP_010809333.1	
WP_010811059.1	WP_010810335.1	WP_010809327.1	
WP_010811057.1	WP_010810334.1	WP_010809326.1	
WP_010811056.1	WP_010810328.1	WP_010809324.1	
WP_010811055.1	WP_010810327.1	WP_010809323.1	
WP_010811054.1	WP_010810326.1	WP_010809321.1	
WP_010811052.1	WP_010810325.1	WP_010809320.1	
WP_010811050.1	WP_010810324.1	WP_010809318.1	
WP_010811049.1	WP_010810321.1	WP_010809301.1	
WP_010811044.1	WP_010810317.1	WP_010809300.1	
WP_010811042.1	WP_010810316.1	WP_010809298.1	
WP_010811037.1	WP_010810315.1	WP_010809297.1	
WP_010811027.1	WP_010810313.1	WP_010809296.1	
WP_010811018.1	WP_010810312.1	WP_010809295.1	
WP_010811007.1	WP_010810310.1	WP_010809292.1	
WP_010811006.1	WP_010810309.1	WP_010809290.1	
WP_010811005.1	WP_010810303.1	WP_010809289.1	
WP_010811004.1	WP_010810302.1	WP_010809288.1	
WP_010811003.1	WP_010810301.1	WP_010809285.1	
WP_010811002.1	WP_010810299.1	WP_010809284.1	
WP_010811001.1	WP_010810296.1	WP_010809281.1	
WP_010810999.1	WP_010810292.1	WP_010809274.1	
WP_010810998.1	WP_010810288.1	WP_010809273.1	
WP_010810995.1	WP_010810286.1	WP_010809270.1	
WP_010810994.1	WP_010810285.1	WP_010809269.1	
WP_010810992.1	WP_010810284.1	WP_010809267.1	

*The green cell presented the type 1 BVMO encoding regions.

Table 44. The nucleotide encoding regions in megaplasmid pHG1 of *C.necator* H16.

WP_193385853.1	WP_011154212.1	WP_041688635.1	WP_011154059.1
WP_193385852.1	WP_011154211.1	WP_041688629.1	WP_011154057.1
WP_193385851.1	WP_011154209.1	WP_011154339.1	WP_011154055.1
WP_162166401.1	WP_011154208.1	WP_011154338.1	WP_011154051.1
WP_162166400.1	WP_011154204.1	WP_011154337.1	WP_011154050.1
WP_157886317.1	WP_011154203.1	WP_011154336.1	WP_011154048.1
WP_157886316.1	WP_011154202.1	WP_011154335.1	WP_011154046.1

Appendix B

WP_136227945.1	WP_011154201.1	WP_011154334.1	WP_011154045.1
WP_136227944.1	WP_011154199.1	WP_011154333.1	WP_011154044.1
WP_136227943.1	WP_011154197.1	WP_011154332.1	WP_011154041.1
WP_136227942.1	WP_011154193.1	WP_011154331.1	WP_011154038.1
WP_136227941.1	WP_011154187.1	WP_011154330.1	WP_011154037.1
WP_136227940.1	WP_011154186.1	WP_011154328.1	WP_011154036.1
WP_136227939.1	WP_011154185.1	WP_011154327.1	WP_011154035.1
WP_136227938.1	WP_011154184.1	WP_011154326.1	WP_011154034.1
WP_136227936.1	WP_011154183.1	WP_011154324.1	WP_011154032.1
WP_136227935.1	WP_011154182.1	WP_011154321.1	WP_011154031.1
WP_136227934.1	WP_011154181.1	WP_011154320.1	WP_011154030.1
WP_136227933.1	WP_011154180.1	WP_011154319.1	WP_011154028.1
WP_136227932.1	WP_011154179.1	WP_011154317.1	WP_011154027.1
WP_136227931.1	WP_011154178.1	WP_011154316.1	WP_011154026.1
WP_136227930.1	WP_011154177.1	WP_011154315.1	WP_011154024.1
WP_136227928.1	WP_011154176.1	WP_011154313.1	WP_011154020.1
WP_136227927.1	WP_011154175.1	WP_011154312.1	WP_011154019.1
WP_082236133.1	WP_011154174.1	WP_011154311.1	WP_011154018.1
WP_082236132.1	WP_011154173.1	WP_011154310.1	WP_011154017.1
WP_082236131.1	WP_011154171.1	WP_011154309.1	WP_011154016.1
WP_082236129.1	WP_011154170.1	WP_011154308.1	WP_011154015.1
WP_082236128.1	WP_011154167.1	WP_011154307.1	WP_011154014.1
WP_082236127.1	WP_011154166.1	WP_011154306.1	WP_011154013.1
WP_082236124.1	WP_011154165.1	WP_011154304.1	WP_011154012.1
WP_082236122.1	WP_011154164.1	WP_011154302.1	WP_011154011.1
WP_082236121.1	WP_011154163.1	WP_011154301.1	WP_011154010.1
WP_082236120.1	WP_011154161.1	WP_011154300.1	WP_011154000.1
WP_082236118.1	WP_011154160.1	WP_011154299.1	WP_011153999.1
WP_082236116.1	WP_011154159.1	WP_011154298.1	WP_011153998.1
WP_082236115.1	WP_011154157.1	WP_011154297.1	WP_011153997.1
WP_082236114.1	WP_011154156.1	WP_011154296.1	WP_011153996.1
WP_063834075.1	WP_011154155.1	WP_011154295.1	WP_011153995.1
WP_063834073.1	WP_011154154.1	WP_011154293.1	WP_011153992.1
WP_051398641.1	WP_011154153.1	WP_011154292.1	WP_011153991.1
WP_051398636.1	WP_011154152.1	WP_011154291.1	WP_011153990.1
WP_051398635.1	WP_011154151.1	WP_011154290.1	WP_011153989.1
WP_051398625.1	WP_011154150.1	WP_011154289.1	WP_011153988.1
WP_051398622.1	WP_011154149.1	WP_011154287.1	WP_011153987.1
WP_051398620.1	WP_011154148.1	WP_011154286.1	WP_011153986.1
WP_051398617.1	WP_011154146.1	WP_011154285.1	WP_011153985.1
WP_041688875.1	WP_011154143.1	WP_011154283.1	WP_011153980.1
WP_041688871.1	WP_011154139.1	WP_011154282.1	WP_011153979.1

Appendix B

WP_041688863.1	WP_011154137.1	WP_011154281.1	WP_011153978.1
WP_041688853.1	WP_011154136.1	WP_011154280.1	WP_011153976.1
WP_041688844.1	WP_011154135.1	WP_011154279.1	WP_011153975.1
WP_041688838.1	WP_011154134.1	WP_011154278.1	WP_011153974.1
WP_041688836.1	WP_011154133.1	WP_011154277.1	WP_011153970.1
WP_041688828.1	WP_011154132.1	WP_011154275.1	WP_011153968.1
WP_041688820.1	WP_011154131.1	WP_011154274.1	WP_011153967.1
WP_041688805.1	WP_011154130.1	WP_011154272.1	WP_011153966.1
WP_041688770.1	WP_011154128.1	WP_011154271.1	WP_011153965.1
WP_041688768.1	WP_011154127.1	WP_011154267.1	WP_011153963.1
WP_041688766.1	WP_011154126.1	WP_011154265.1	WP_011153962.1
WP_041688760.1	WP_011154125.1	WP_011154264.1	WP_011153961.1
WP_041688756.1	WP_011154121.1	WP_011154261.1	WP_011153960.1
WP_041688754.1	WP_011154120.1	WP_011154260.1	WP_011153959.1
WP_041688752.1	WP_011154118.1	WP_011154259.1	WP_011153948.1
WP_041688750.1	WP_011154117.1	WP_011154258.1	WP_011153947.1
WP_041688748.1	WP_011154115.1	WP_011154257.1	WP_011153946.1
WP_041688746.1	WP_011154112.1	WP_011154256.1	WP_011153945.1
WP_041688744.1	WP_011154111.1	WP_011154255.1	WP_011153944.1
WP_041688742.1	WP_011154109.1	WP_011154253.1	WP_011153943.1
WP_041688740.1	WP_011154108.1	WP_011154252.1	WP_011153942.1
WP_041688738.1	WP_011154107.1	WP_011154251.1	WP_011153941.1
WP_041688736.1	WP_011154106.1	WP_011154250.1	WP_011153940.1
WP_041688734.1	WP_011154103.1	WP_011154249.1	WP_011153939.1
WP_041688732.1	WP_011154102.1	WP_011154248.1	WP_011153938.1
WP_041688728.1	WP_011154101.1	WP_011154247.1	WP_011153937.1
WP_041688726.1	WP_011154099.1	WP_011154245.1	WP_011153936.1
WP_041688724.1	WP_011154098.1	WP_011154242.1	WP_011153935.1
WP_041688718.1	WP_011154097.1	WP_011154241.1	WP_011153934.1
WP_041688716.1	WP_011154096.1	WP_011154238.1	WP_011153933.1
WP_041688714.1	WP_011154095.1	WP_011154237.1	WP_011153932.1
WP_041688712.1	WP_011154094.1	WP_011154236.1	WP_011153931.1
WP_041688704.1	WP_011154092.1	WP_011154235.1	WP_011153927.1
WP_041688696.1	WP_011154091.1	WP_011154234.1	WP_011153926.1
WP_041688694.1	WP_011154089.1	WP_011154233.1	
WP_041688692.1	WP_011154088.1	WP_011154231.1	
WP_041688690.1	WP_011154086.1	WP_011154229.1	
WP_041688683.1	WP_011154085.1	WP_011154228.1	
WP_041688681.1	WP_011154084.1	WP_011154227.1	
WP_041688677.1	WP_011154083.1	WP_011154226.1	
WP_041688675.1	WP_011154082.1	WP_011154225.1	
WP_041688671.1	WP_011154079.1	WP_011154224.1	

Appendix B

WP_041688667.1	WP_011154078.1	WP_011154223.1	
WP_041688660.1	WP_011154077.1	WP_011154222.1	
WP_041688658.1	WP_011154075.1	WP_011154221.1	
WP_041688654.1	WP_011154074.1	WP_011154220.1	
WP_041688652.1	WP_011154072.1	WP_011154219.1	
WP_041688647.1	WP_011154066.1	WP_011154218.1	
WP_041688645.1	WP_011154064.1	WP_011154216.1	
WP_041688643.1	WP_011154062.1	WP_011154214.1	
WP_041688639.1	WP_011154060.1	WP_011154213.1	

Appendix C Abbreviations

Table 45. Abbreviations of amino acids.

Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Table 46. General Abbreviations.

(v/v)	Volume per volume	FAH	Fatty acid hydratases
(w/v)	Weight per volume	fwd	Forward
(w/w)	Weight per weight	g	Earth's gravitational field
°C	Degrees Celsius	GFP	Green fluorescent protein
ADO	Aldehyde deformylating oxygenase	GMC	Glucose-methanol choline
ASR	Ancestral sequence reconstruction	GC	Gas chromatography
BLAST	Basic Local Alignment Search Tool	GC-FID	Gas chromatography with flame ionization detection
bp	Base pair	GC-MS	Gas chromatography with mass spectrometry
BVMO	Baeyer-Villiger Monooxygenases	gDNA	Genomic DNA
BSA	Bovine serum albumin	h	Hours
CO₂	carbon dioxide	IPTG	Isopropyl β-D-1-thiogalactopyranoside
CFE	Cell-free extract	Kan	Kanamycin
ChCl	Choline chloride	kb	Kilobase (10 ³ bases)
cPCR	Colony polymerase chain reaction	kDa	Kilodalton (10 ³ Da)
CV	Column volume	LPG	Liquefied Petroleum Gas
CWW	Cell wet weight	M	Molar (mol L ⁻¹)
DCM	Dichloromethane	NADPH	Nicotinamide adenine dinucleotide phosphate
DMSO	Dimethylsulfoxide	NADH	Nicotinamide adenine dinucleotide
ddH₂O	Double-distilled water	min	Minute
DNA	Deoxyribonucleic acid	mL	Milliliter (10 ⁻³ L)
DES	Deep eutectic solvent	mm	Millimeter (10 ⁻³ m)
DME	Dimethyl ether	mM	Millimolar (mmol L ⁻¹)
dH₂O	Distilled water	n.d.	Not determined
dNTPs	Deoxynucleotide Triphosphates	ng	Nanogram (10 ⁻⁹ g)
EC	Enzyme Commission	nm	Nanometer (10 ⁻⁹ m)
ee	Enantiomeric excess	rpm	Revolutions per minute
EtOAc	Ethyl acetate	SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
FAD	Flavin adenine dinucleotide	SSM	Site-saturation mutagenesis
FAP	Fatty acid photodecarboxylase	T	Temperature (in °C)
FAS	Fatty acid synthesis	TrxA	Thioredoxin A

Appendix C

OleABCD	Thiolase-based system	t	Time
Ols	Olefin synthase	Tm	Melting temperature (in °C)
OleT_{JE}	Olefin-forming fatty acid decarboxylase	U	Enzyme unit
OD600	Optical density at 600 nm	UV	Ultraviolet light
ONC	Overnight culture	VFA	Volatile fatty acid
PPT	Phosphinothricin	V	Volt
PCR	Polymerase chain reaction	var.	Variable
PDB	Protein data bank	WT	Wildtype
rev	Reverse	A	Angstrom (10-10 m)

Appendix D List of Figures

Figure 1. Classification of biofuels.	5
Figure 2. The working mechanism of microalgae in the energy sector.	6
Figure 3. Contribution of bio-based hydrocarbon from low-value substrates via metabolic engineering.	8
Figure 4. Overall diagram of enzymatic reaction for converting precursors substrates to alka(e)nes.	8
Figure 5. Recently published study about CvFAP.	11
Figure 6. Suggested CvFAP photocycle.	12
Figure 7. Photoenzymatic cascade to transform triolein into (Z)-heptadec-8-ene.	13
Figure 8. Overall scheme of the biocatalytic production of gaseous hydrocarbons from amino acids.	14
Figure 9. Photoenzymatic cascades to transform unsaturated fatty acid into secondary alcohols.	14
Figure 10. Asymmetric synthesis of α -hydroxy/amino acids by CvFAP.	15
Figure 11. Enzymatic asymmetric synthesis of L-phosphinothricin (L-PPT).	15
Figure 12. Metabolic pathways for alka(e)ne synthesis in <i>Yarrowia lipolytica</i>	16
Figure 13. Phylogenetic relationship of the 12 sequences of extant FAP enzymes used to reconstruct the ancestral sequences.	19
Figure 14. Pairwise identity between each ancestor and CvFAP.	20
Figure 15. Multiple sequence alignment of CvFAP, ANC1, ANC2, and ANC3 using Clustal Omega.	21
Figure 16. Homology models of ancestors superimposed on CvFAP structure (PDB-ID: 5NCC).	22
Figure 17. SDS-PAGE analysis of different fractions from Ni-affinity purification of CvFAP, ANC1, ANC2 and ANC3.	23
Figure 18. SDS-PAGE analysis of different fractions from Ni-affinity purification of sANC1 and TrxA-sANC1.	24
Figure 19. Comparison between CvFAP, ANC1, and sANC1 produced using two different IPTG concentrations.	26
Figure 20. Comparison of sCvFAP, ANC1, and sANC1 residual activities after being exposed to light at 455 nm for the stated amount of time.	27
Figure 21. Comparison of sCvFAP, ANC1, and sANC1 product production at different external FAD concentrations.	27
Figure 22. Time-resolved emission spectroscopy for different concentrations of CvFAP.	28
Figure 23. Time-resolved emission spectroscopy of CvFAP, ANC1.	29
Figure 24. Proposed photocycle of CvFAP.	30
Figure 25. First derivative curves of the melting temperature curves obtained in fluorescence measurements.	31
Figure 26. Thermostability of CvFAP, ANC1, and SANC1 produced using different IPTG concentrations (0.1 mM in A and 0.5 mM in B).	33
Figure 27. ProtScale hydrophobicity graph for CvFAP (pink) and ANC1 (orange).	34

Figure 28. Flexibility comparison of CvFAP and ANC1.....	35
Figure 29. Time course of pentadecane production of lyophilized whole <i>E.coli</i> BL21(DE3) expressing CvFAP or ANC1.	36
Figure 30. First derivative curves of the melting temperature curves obtained in fluorescence measurements for CvFAP and ANC1 in different DESs concentrations...	41
Figure 31. Time course of pentadecane production of CvFAP and ANC1 in DESs.	42
Figure 32. Hydrophobic DES synthesized in this study	42
Figure 33. The stable hits from Hotspot Wizard prediction (Top); Point mutant (pmut) scanning score made via Rosetta scoring.	45
Figure 34. SDS-PAGE analysis of CvFAP mutants.	46
Figure 35. Production of pentadecane with CvFAP and its mutants (left) and thermostability of CvFAP and mutants (right).	46
Figure 36. ThermoFAD experiment for CvFAP WT and mutant G152N, S473G.	47
Figure 37. Sequence alignment of CvFAP and ANC1.	48
Figure 38. Structure of active site of sANC1 and CvFAP with palmitic acid.....	48
Figure 39. SDS-PAGE analysis of sANC1 and variants (MW=66 kDa) expression.	49
Figure 40. Specific activity of sANC1 mutants	49
Figure 41. Chemical structure of Nile red. The dye is hydrophobic and fluorescent.	50
Figure 42. Flowchart of the 96-well plate format high-throughput screening strategy via Nile red.	51
Figure 43. Emission spectra of Nile red-stained pentadecane and palmitic acid.	51
Figure 44. Emission spectra of Nile red-stained pentadecane and palmitic acid.	52
Figure 45. Fluorescence emission spectra of Nile red-stained cell-free extract and whole-cell of FAPs.	53
Figure 46. Thermodynamic cycle in Rossta energy calculation.	73
Figure 47. Workflow of HotSpot Wizard 3.0.	74
Figure 48. Interface of the prediction result output in the HotSpotWizard 3.0.	74
Figure 49. Classification of oxygenases and enzymatic oxidation reactions as catalyzed by monooxygenases and dioxygenases.	81
Figure 50. Oxidation of menthone by Baeyer and Villiger. Figure adapted from Leisch. ^[151]	83
Figure 51. Baeyer–Villiger oxidation and sulfoxidation reactions catalyzed by flavin-dependent Baeyer–Villiger monooxygenases (BVMOs).....	84
Figure 52. Taxonomic distribution of BVMO genes. ^[176]	86
Figure 53. The distribution of the identified BVMO genes on <i>C.necator</i> H16 chromosome 1 and 2.....	90
Figure 54. Multiple nucleotide sequence alignment of CPMO0, CPMO11, and CPMO13 from <i>C.necator</i> H16 with known BVMOs.....	92
Figure 55. The whole cell-mediated biotransformation from Cyclopentanone to 5-Valerolactone.....	95
Figure 56. Expression analysis of CPMO0.	96
Figure 57. SDS-PAGE analysis of CPMO0 solubilization.	97
Figure 58. SDS-PAGE analysis co-expression CPMO0 with chaperon protein in <i>E.coli</i>	98

Figure 59. SDS-PAGE analysis of the expression of CPMO0.	99
Figure 60. SDS-PAGE analysis of the expression of CPMO0 with soluble tag: TrxA in <i>E.coli</i>	100
Figure 61. SDS-PAGE analysis of the expression of CPMO0 in <i>C.necator</i> H16.....	102
Figure 62. GC-FID Chromatogram of substrates and products.	103
Figure 63. First derivative curves of the melting temperature curves obtained in fluorescence measurements.....	104
Figure 64. Multiple nucleotide sequence alignment of CPMO1 from <i>C.necator</i> H16 with known BVMOs.....	165
Figure 65. Multiple nucleotide sequence alignment of CPMO2 from <i>C.necator</i> H16 with known BVMOs.....	165
Figure 66. Multiple nucleotide sequence alignment of CPMO3 from <i>C.necator</i> H16 with known BVMOs.....	166
Figure 67. Multiple nucleotide sequence alignment of CPMO4 from <i>C.necator</i> H16 with known BVMOs.....	166
Figure 68. Multiple nucleotide sequence alignment of CPMO5 from <i>C.necator</i> H16 with known BVMOs.....	167
Figure 69. Multiple nucleotide sequence alignment of CPMO6 from <i>C.necator</i> H16 with known BVMOs.....	167
Figure 70. Multiple nucleotide sequence alignment of CPMO7 from <i>C.necator</i> H16 with known BVMOs.....	168
Figure 71. Multiple nucleotide sequence alignment of CPMO8 from <i>C.necator</i> H16 with known BVMOs.....	168
Figure 72. Multiple nucleotide sequence alignment of CPMO9 from <i>C.necator</i> H16 with known BVMOs.....	169
Figure 73. Multiple nucleotide sequence alignment of CPMO10 from <i>C.necator</i> H16 with known BVMOs.....	169
Figure 74. Multiple nucleotide sequence alignment of CPMO11 from <i>C.necator</i> H16 with known BVMOs.....	170
Figure 75. Multiple nucleotide sequence alignment of CPMO12 from <i>C.necator</i> H16 with known BVMOs.....	170
Figure 76. Multiple nucleotide sequence alignment of CPMO13 from <i>C.necator</i> H16 with known BVMOs.....	171
Figure 77. Multiple nucleotide sequence alignment of CPMO14 from <i>C.necator</i> H16 with known BVMOs.....	171
Figure 78. Multiple nucleotide sequence alignment of CPMO15 from <i>C.necator</i> H16 with known BVMOs.....	172
Figure 79. Multiple nucleotide sequence alignment of CPMO16 from <i>C.necator</i> H16 with known BVMOs.....	172
Figure 80. Multiple nucleotide sequence alignment of CPMO17 from <i>C.necator</i> H16 with known BVMOs.....	173

Appendix E List of Tables

Table 1. Conventional transportation fuel properties and potential biofuels	7
Table 2. Purification yield of CvFAP and ancestor enzymes.	23
Table 3. Purification yields of CvFAP, ANC1, and sANC1 enzyme.	25
Table 4. Comparison of T _m of both peaks determined by ThermoFAD for CvFAP, ANC1, and sANC1 in aqueous buffer with or without 30% DMSO.	32
Table 5. Decarboxylation of ANC1 and sANC1 of various fatty acid substrates compared to wild-type CvFAP.	37
Table 6. Conversion of palmitic acid to pentadecane using different mixtures of DES and water	38
Table 7. Ingredients concentration of DESs	39
Table 8. Comparison of T _m of both peaks determined by ThermoFAD for CvFAP, ANC1, in Tris-HCl buffer with different DES components.	40
Table 9. Synthesis of Hydrophobic DES synthesized in this study.....	43
Table 10. Production of Heptane and Nonane in DES	44
Table 11. Summary of predicted CvFAP variants	45
Table 12. DNA sequences used in this study.....	57
Table 13. Amino acid sequence used in this study	61
Table 14. Recombinant plasmids used in this study.	65
Table 15. Primers used in this study. Mutated bases are shown in bold letters.	66
Table 16. Sample composition (20 μL) and temperature program for colony PCR using a ready-to-use 2x PCR Master Mix.....	68
Table 17. Sample composition (20 μL) and temperature program for QuikChange reaction.	68
Table 18. Method parameters used for GC-FID analysis.	72
Table 19. Chemicals used in this study.....	75
Table 20. Composition of general media, buffers, and solutions used in this study.	75
Table 21. Buffers and solutions for chemo-competent <i>E.coli</i> cells preparation in this study.	76
Table 22. Buffers for enzyme purification in this study.	76
Table 23. Buffers and solutions for SDS-PAGE in this study.	77
Table 24. Flavin monooxygenase classification. ^[149]	82
Table 25. General Characteristics and Classification of BVMOs.....	84
Table 26. General features of the <i>Cupriavidus necator</i> H16 genome.....	87
Table 27. Overview of putative BVMOs encoding gene from <i>C.necator</i> H16.....	89
Table 28. Sequence motifs of putative BVMOs from <i>C.necator</i> H16.....	91
Table 29. Pairwise sequence identities between CPMO0 and CPMO11 from <i>C.necator</i> H16 with typical BVMOs.	93
Table 30. Substrates scope screening of CPMO0.	94
Table 31. Chaperone Plasmid pGro7	98
Table 32. Comparison of T _m determined by ThermoFAD for CPMO0.....	104
Table 33. Strains used in this study	107
Table 34. Plasmids used in this study	107

Table 35. List of primers used in this study.	108
Table 36. Components of the FastCloning PCR reactions.	110
Table 37. Master mixture for Gibson Assembly.....	111
Table 38. 5X ISO buffer for Gibson Assembly.....	111
Table 39. Method parameters used for GC-FID analysis.	116
Table 40. Retention times (given in minutes) of identified compounds relevant to this work	116
Table 41. The amino acid sequence used in this study.....	174
Table 42. The nucleotide encoding regions in chromosome 1 of <i>C.necator</i> H16.	156
Table 43. The nucleotide encoding regions in chromosome 2 of <i>C.necator</i> H16.	175
Table 44. The nucleotide encoding regions in megaplasmid pHG1 of <i>C.necator</i> H16.	189
Table 45. Abbreviations of amino acids.....	193
Table 46. General Abbreviations.	194