Cytochrome P450 whole cell biohydroxylation of alkanes

Danielle Seeger

In fulfilment of the requirements for the degree of Master of Science in Engineering

Supervisor: **Prof Sue Harrison**
Co-supervisor: **Dr Athanasios Kotsiopoulos**

Centre for Bioprocess Engineering Research (CeBER)

Department of Chemical Engineering

Faculty of Engineering and the Built Environment

University of Cape Town

July 2023

July 23

Plagiarism Declaration

1. I know that plagiarism is wrong. Plagiarism is to use another's work and to pretend that it is ones own.

2. I have used the Harvard system for citation and referencing. Each significant contribution to, and quotation in, this report from the work, or works, of other people has been attributed, and has been cited and referenced.

3. This report is my own unaided work, except for assistance received from the teaching staff.

4. I have not allowed and will not allow anyone to copy my work with the intention of passing it off as his or her own work.

Signed by candidate

_____

Signature

Abstract

Production of higher-value fine chemicals via bio-based catalytic alkane activation are becoming increasingly relevant in the context of sustainability. Linear alkanes, sourced from fossil fuel reserves and as by-products in gas-to-liquid technology and CO_2-to-power strategies, both through Fischer Tropsch activities, can serve not only as a low-cost fuel or solvent, but also as a precursor to produce valuable long chain alcohols. Effective use, or resource efficiency, with respect to alkanes is important in sustainable processing. Alkanes contain unreactive hydrocarbon bonds and cracking processes apply high temperatures and pressures to split these inert molecules into smaller and reactive molecules. Although widely applied, cracking suffers from poor selectivity and is resource heavy and environmentally burdening. Biocatalysis has shown potential for highly regio-selective alkane oxidation under mild conditions. Enzyme catalytic mechanisms evolved to catabolise fatty molecules in an organism's metabolic pathway. Biocatalysis harnesses these evolved pathways to produce value-added chemicals.

Cytochrome P450 (CYP) monooxygenases such as CYP153A6 have previously shown excellent 95% selectivity for hydroxylation at the terminal carbon on *n*-octane and CYP153A13 has shown promising activity on *n*-octane and *n*-decane. In the present study, whole cell biocatalysts were prepared by expressing CYPs in *E. coli* BL21(DE3) and suspended in aqueous media required to maintain whole cell health. Alkanes applied were immiscible with the aqueous phase, resulting in a '2-Liquid Phase System' (2LPS). Alkane accessibility to the enzyme in the whole cells significantly constrains scalability of biocatalysis due to transfer across the aqueous-organic interface and the cellular membrane to reach CYPs. The extent to which alkanes are transferred is determined by variables such as solubility, thermodynamic equilibrium, cell membrane permeability, temperature etc. The kinetics of the biocatalytic reaction are partly dependent on physicochemical variables, but also on biological parameters such as cellular health, enzyme expression etc.

It was shown that although CYP153A13 showed superior coupling efficiencies compared to multicomponent enzyme systems, fusion proteins suffer from low expression levels and cytoplasmic proteolytic truncation, lowering its concentration and activity compared to CYP153A6. High Cell Density Cultures (HCDCs) were used to circumvent CYP153A13's low expression. HCDCs of both CYP153A6 and CYP153A13 resulted in larger 1-octanol volumetric production rates by 10- and four-fold respectively compared to the LCDC counterparts. The turnover number of CYP153A6 and -A13 LCDCs, however, were respectively 2- and 7- fold higher than their HCDC counterparts, illustrating the oxygen and nutrient limitation of HCDCs.

The multicomponent CYP153A6 system was expressed with high functionality and showed descending activity with: *n*-octane > *n*-decane > *n*-nonane, caused by a combination of alkane solubility and transmembrane transport. Although *n*-decane is less soluble in water than *n*-nonane, the 'odd-even' effect of alkanes resulted in high *in vivo* activity on *n*-decane. UNIF-LL predictions illustrated that BEHP reduces the amount of alkane in the aqueous phase than in the absence of BEHP. Therefore, it was illustrated that BEHP regulates the amount of substrate in the aqueous phase, which decreases the toxicity of the alkane substrates towards the whole cells. Therefore, while substrate solubility plays a role in availability to the whole cell biocatalyst which would impact the rate of biocatalytic reaction, cellular health maintenance is also important.

This proved that factors related to aqueous substrate solubility play a larger role in the outcomes of 2LPS whole cell biocatalytic reaction kinetics as well as factors such as cellular health. The study showed that although substrate solubility plays an important role in biocatalytic performance, substrate

solubility could not be used as the sole indicator of predicting biocatalytic performance and factors such as enzyme expression, membrane transport, enzyme-substrate affinity and physical substrate properties should be considered. To investigate the enzyme-substrate affinity, apparent Michaelis-Menten (MM) constants (K_m^c) of each enzyme-substrate pair were determined from experimental data, of product concentration, accumulated over time using least squares nonlinear regression.

K_m^c constants, however, represented both enzyme-substrate (ES) and substrate-solvent interactions. To determine the true affinity of each ES pair without the interference of substrate-solvent interactions on ES interactions, thermodynamic activity-based constants (K_m^a) were determined. This approach paves the way for the future development of 'one-pot' biocatalytic systems containing multiple different enzymes and substrates. A process flowsheet using the UNIF-LL model was developed and a theoretical calculation were used to determine activity coefficients, which were in turn used to determine the K_m^a constants. The K_m^c constants of CYP153A6 illustrated a 17-fold and two-fold larger affinity for *n*-decane and *n*-nonane compared to *n*-octane, respectively. By excluding the effects of substrate-solvent interactions, CYP153A6 had an affinity twice as large for *n*-nonane compared to both *n*-octane and *n*-decane. Therefore, while the K_m^c constants depicted that *n*-decane had the highest affinity with CYP153A6, this resulted from unfavourable molecular interactions between *n*-decane and the aqueous solvent, rather than favourable ES affinity. The K_m^c constants suggested that *n*-decane is the preferred substrate of CYP153A13 by two-fold compared to *n*-octane, but K_m^a values revealed that CYP153A13 had a six-fold larger affinity for *n*-octane than *n*-decane.

The UNIF-LL model predicted substrate available in the aqueous phase at phase equilibrium for each alkane in the 2LPS. These predicted aqueous phase substrate concentrations were compared to the enzyme concentration and the MM constants, respectively. From these comparisons, it was determined that the reactions catalysed by CYP153A6 on *n*-octane and *n*-nonane were not limited by substrate concentration, whilst the reactions catalysed by CYP153A13 on *n*-decane were limited by substrate concentration. These results, together with enzyme expression characterisation mentioned earlier, means that CYP153A6 processes can be improved using enzyme engineering techniques, whilst CYP153A13 processes can be improved using enzyme engineering to improve enzyme expression levels as well as substrate delivery enhancement techniques.

These findings highlighted the impact of substrate solubility on substrate partitioning into the aqueous phase to be available to the biocatalyst. However, once the substrate reaches the aqueous phase, effective bioconversion relies on adequate transmembrane transport, sufficient enzyme expression levels, ES affinity, and the effectiveness of the reaction mechanism. Therefore, both physiochemical and biological factors play a role in the demand-based delivery system. This demand-based delivery system is governed by the rate of substrate transfer across the liquid-liquid and membrane barriers, as well as the amount of substrate converted into product by the catalytic mechanism of the enzyme, thereby allowing more substrate to partition into the aqueous phase.

Overall, the approach developed could be used to enhance the kinetic studies of CYP153 whole cell alkane primary hydroxylation in two-liquid-phase-systems, a field that receives little attention but shows importance in the goal of scaling up such systems. Globally, researchers and industry share interest in the biohydroxylation of linear chain alkanes due to their high-end properties and value, but poor solubility limits techno-economic feasibility and large-scale application. This study showed promising compatibility of CYP enzymes with medium chain alkanes (C_8 C_9 C_{10}), but showed the challenges presented in molecular biology such as enzyme expression. Moreover, substrate solubility of linear alkanes impacts the location of the reaction within a 2LPS, accessibility and the apparent ES affinity.

Overall, the study developed an integrated approach to investigate the reaction kinetics of an alkane whole cell hydroxylation 2LPS by considering the limitations arising from both molecular biology and biochemistry as well as thermodynamics. From this, the interdisciplinarity of biocatalysis was enhanced by integrating aspects of molecular biology and biochemistry into the greater context of chemical engineering. The findings can be used to facilitate research in alkane bio-activation process control, optimisation and scaling up by implementing the use of multiple enzymes and substrates simultaneously.

Acknowledgements

I wish to extend my sincere gratitude to the following people and institutions for their contribution to this research project:

My supervisor Professor Sue Harrison and my co-supervisor Doctor Athanasios Kotsiopoulos for their wisdom, assistance, and guidance throughout my research.

Doctor Mariette Smart for assisting me with laboratory work. Shanna Swart for always being reliable and helpful when I asked, assisting me with laboratory work and providing me with insights into the research field of biocatalysis.

Thank you to Tich Samkange and Sharon Rademeyer for always being helpful in the laboratory.

My family - my mom and dad, Celeste and Erich Seeger for their continued faith in my abilities, and support throughout the years of my degree. To my grandparents Annemarie and William Stillwell for their continuous support love and encouragement. I'd like to give a special thank you to my grandpa (my boka!) for initially inspiring me to study genetics.

The CeBER team at the University of Cape Town for all their support and lovely chats in the tearoom and in the laboratory!

Table of Contents

List of Figures

List of Tables

Glossary of Terms

a	Thermodynamic activity
Biocatalyst efficiency	Efficiency measured in mass of product per mass of dry cell weight whole cell biocatalyst
c_{alkane}^{aq}	alkane concentration in aqueous phase
$c_{alkane\ UNIF-LL}^{aq}$	UNIF-LL predicted alkane concentration in aqueous phase
Induction	The process of inducing gene expression to allow the formation of recombinant protein
k_{cat}	Catalytic efficiency
K_D	Binding/dissociation constant
K_m	Michaelis-Menten constant
K_m^c	Substrate-concentration-based Apparent Michaelis-Menten constant inclusive of enzyme-substrate interactions and solvent-substrate interactions
K_m^a	Thermodynamic activity-based Michaelis-Menten constant inclusive of only enzyme-substrate interactions
K_p	Distribution ratio in terms of species concentration in a two-liquid-phase-system
Liquid-Liquid	Two immiscible liquids on contact with one another
Maximum specific biocatalytic activity	Biocatalytic activity measured in mole of product per dry cell weight per unit time
Maximum reaction velocity	The maximum velocity of a reaction in which saturation is reached i.e., all enzyme active sites are occupied with substrate using the Michaelis-Menten model
Michaelis-Menten constant	The substrate concentration at which half the maximum velocity is reached of a reaction also known as K_m
OD_{600}	Optical density at a wavelength of 600nm
P_x	Partition coefficient
Specific intracellular P450 concentration	The P450 concentration measured in mole of P450 per mass of dry cell weight

t_0	Reaction start time
Turnover number (TON)	The turnover number (TON), indicating the mole of product per mole of active enzyme, also known as specific product concentration under defined reaction conditions. The turnover number is also known as the first order rate constant (k_{cat}) to indicate the number of substrate molecules bioconverted per unit time when the enzyme is at maximum efficiency.
Turnover frequency (TOF)	The turnover frequency (TOF), indicating the mole product produced per mole of active enzyme to catalyse the reaction site per unit time, also known as specific activity under defined reaction conditions
v_0	Initial reaction velocity
v_{max}	Maximum reaction velocity
Volumetric production rate	Productivity of a biocatalysed reaction measured in mole of product per working volume per unit time
x_{alkane}^{org}	Experimental molar fraction of alkane in organic phase
$x_{alkane\ UNIF-LL}^{aq}$	Predicted molar fraction of alkane in the aqueous phase
γ	Thermodynamic activity coefficient

Acronyms and Abbreviations

2-Liquid-Phase-System	2LPS
5-ala	5-aminolevulinic acid
AlkB	Alkane-1-monoxygenase
BEHP	Bis(2-ethylhexyl) phthalate
BRM	Bioreaction Mixture
CYP	Cytochrome P450 monooxygenase
DCW	Dry cell weight
dH_2O	De-ionised water
DMSO	Dimethyl sulfoxide
ES	Enzyme-Substrate
EtAc	Ethyl acetate
FdR	Ferredoxin Reductase
FdX	Ferredoxin
GC	Gas chromatography
GCM	Group Contribution Method
GTL	Gas-To-Liquid
IPTG	Isopropyl-β-D-thiogalactopyranoside
LL	Liquid-Liquid
LB	Luria Bertani
MM	Michaelis-Menten
NAD(P)H	nicotinamide adenine dinucleotide phosphate
NLR	Nonlinear regression
rpm	Revolutions per minute
UNIF-LL	Universal quasichemical Functional group Activity Coefficients Liquid-Liquid model

1 Alkane activation

1.1 Significance and challenges of current alkane activation methods

Alkanes are saturated hydrocarbon molecules that are abundantly available, inexpensive, and are by-products of expanding petrochemical operations such as gas-to-liquid processes. Alkanes are highly inert and stable due to their strong C-H covalent bonds (Labinger and Bercaw, 2002). Alkanes' most economically profitable trait is their potential to serve as precursor molecules for higher-value platform and fine chemicals production. Platform and fine chemical production require activated intermediates such as fatty alcohols, aldehydes, and carboxylic acids. The hydrocarbon backbone can be used as a basic structure upon which more reactive functional groups can be added to produce these activated intermediates. The value and market relevance of linear saturated alkanes is increased when reactive functional groups are added to the hydrocarbon backbone.

Diesel is a naturally found mixture of paraffinic long-chain hydrocarbons and is typically used as fuel. Fatty alcohols, which differ from diesel by only a single functional hydroxyl group, on the other hand, have broader industrial applications and, therefore, an increased market value (Lennen and Pfleger, 2013). Fatty alcohols are molecules that possess a single hydroxyl group at the terminal carbon of a long straight-chain hydrocarbon backbone. The market demand for fatty alcohol is increasing rapidly and was estimated to be worth € 1 billion in Europe alone in 2011 (Grant et al., 2011). Chain length determines end-product properties, industrial application, and value (Lennen and Pfleger, 2013). Medium to longer-chain alcohols, for example, can be used in detergents, hygiene products, and other industries and are extremely valuable (Grant et al., 2011). Production processes that allow for precise and selective control of product chain length and modification are, therefore, highly desirable (Lennen and Pfleger, 2013).

Alkane functionalisation is of great interest to the global petrochemical and chemical syntheses industry. The challenge of converting this abundantly available and inexpensive feedstock to produce fine chemical intermediates has intrigued scientists for many years. Chemists, petrochemical and pharmaceutical industries have made long-standing global efforts to develop methods to functionalise linear alkanes. Indeed, an attractive feedstock for chemical production, but the most significant challenge remains - linear alkanes possess one of the strongest bonds in nature. The covalent bonds in alkanes possess approximately 400 kJ/mol of energy, which can be harnessed via combustion at high temperatures (Shilov and Shul'pin, 1997).

Chemical cracking is another process that can convert alkanes to higher-value olefin products (Ayala and Torres, 2004). Cracking is a large-scale industrial alkane hydroxylation process. Cracking via chemical catalytic methods utilise harsh oxidants, high pressure, and temperatures up to 800˚C to break these strong covalent bonds, causing random oxidation, activation and functionalisation of alkanes into more useful and reactive hydrocarbon molecules, such as alkenes (Ayala and Torres, 2004). Alkanes can also react when exposed to highly reactive conditions such as free radicals. Alkenes have higher reactivity and are preferred as starting materials for chemical conversion over alkanes, but alkanes are four times cheaper. Therefore, fine chemical production using alkanes as starting material is more economically favourable, provided a suitable process can be determined.

Chemical approaches to alkane activation are energy intensive, resource demanding, environmentally burdening, and often suffer from poor selectivity, over-oxidised products, toxic metal catalysts, and unwanted side reactions, causing increased downstream processing costs and environmental burden

(Shilov and Shul'pin, 1997). Chemo-catalytic oxy-functionalisation routes are susceptible to over-oxidised products and poor selectivity because partially oxidised products are more reactive than their respective reactants (Cavani and Trifiro, 1997). Biocatalysis is an emerging and promising approach for the highly selective oxidation of alkanes under mild conditions. The present study uses whole cell biocatalysis to produce medium-chain primary alcohols from alkane feedstocks.

1.2 Biocatalytic alkane activation

Biocatalysis has become increasingly recognised as an alternative and practical approach to activating alkanes for fine chemical synthesis. Biocatalysis is an emerging alternative approach to highly selective alkane activation under milder process and operation conditions, with reduced side-product formation, highly selective product formation, and compatibility with an extensive range of complex substrates. Biocatalysis has massive potential for the industrial production of complex fine pharmaceuticals, polymers, adhesives, detergents, plasticisers, fragrances, additives, flavours, and agrochemicals. Moreover, bio-based chemical production processes are becoming relevant in the context of universally ratified sustainable development goals.

Biocatalysis utilises the specialised design and targeted mechanisms of enzymes as alternative, practical, and “greener” approaches to complex chemical reactions such as alkane activation. This is mainly due to its ability to catalyse complex and selective reactions at mild operation conditions. Compared to conventional chemical cracking, biocatalytic fine chemical production has the potential to reduce toxic solvent usage, production costs, downstream processing costs, energy consumption, and environmental burden.

The heart of biocatalysis lies within a group of functional proteins named enzymes. Enzymes act as catalysts for biochemical reactions in metabolic pathways. Enzymes have been carefully designed through evolution over millions of years to catabolise fatty molecules in an organism’s natural metabolism. The exponentially growing field of biotechnology augments the industrial potential of biocatalytic activation of alkanes and bio-based fine chemical syntheses. Nature offers a vast selection of genes, alkane oxygenase proteins, and host organisms to facilitate the development of alkane activation processes and biocatalytic tools. Biocatalysts can be crafted to suit the substrate of interest, recombinant host cell, and application. Some alkane monooxygenases include methane monooxygenases, integral membrane non-heme alkane hydroxylases, and cytochrome P450s. The present study employed cytochrome P450 (CYPs) monooxygenases for reasons discussed below.

1.3 Cytochrome P450 monooxygenases for alkane oxidation

The CYP enzyme family are heme-containing enzymes mainly studied for empowering various organisms to metabolise or functionalise aliphatic compounds. Their application ranges from xenobiotic metabolism and biosynthesis of steroids by human cells to the bioremediation of industrial inert chemical pollutants. This enzyme superfamily has gained attention in biocatalysis for its ability to introduce an oxygen molecule into a non-functionalised carbon chain alkane with high regio- and stereo-selectivity. Several members of this enzyme superfamily have been discovered in various alkane-degrading bacterial species, including *Pseudomonas putida., Mycobacterium* sp*., Acinetobacter* sp*,* and *Alcanivorax borkumensis,* amongst many others (Funhoff et al., 2006). The CYPs isolated from these organisms can be recombinantly expressed, engineered, and implemented in alkane activation studies.

CYPs are well-recognised for their ability to selectively oxy-functionalise, under mild operation conditions, the C-H bond found in a wide range of aliphatic compounds, including limonenes, steroids,

styrene, naphthalene, and alkanes. Their broad substrate compatibility paves the way for exploring potential industrially relevant and useful reactions. CYP153s have been notably recognised for their ability to catalyse the hydroxylation of linear *n*-alkanes at the terminal carbon to produce valuable primary alcohols with substantially high selectivity for the terminal position (Kochius et al., 2018). CYP153A6 from *Mycobacterium* sp. strain HXN-1500 is the most well-studied CYP153 family and prefers C_6 to C_{11} alkanes with ≥95% selective oxidation for the terminal carbon (Koch et al., 2009a, Funhoff et al., 2006).

CYPs catalyse the oxidation of aliphatic compounds by concomitant reduction of molecular oxygen. Electrons required for this redox reaction are passed on from redox partner proteins after cofactor nicotinamide adenine dinucleotide phosphate (NAD(P)H) oxidation. The mechanism of the CYP-based biocatalytic alkane oxidation relies on shuttling of electrons from NAD(P)H to the heme prosthetic group (Bordeaux et al., 2011). Electrons, however, can only be transferred from the cofactor to the biocatalytic active site with the assistance of electron transfer from redox partner proteins. Whole cell biocatalysis processes are estimated to be ten times cheaper than purified, crude, or immobilised enzyme processes since they regenerate NAD(P)H and remove expensive enzyme purification processes (Wachtmeister and Rother, 2016).

A biocatalytic reaction's coupling efficiency indicates the correct transfer of electrons to the substrate of interest from the cofactor. Low coupling efficiency and uncoupled electron transfer can affect biocatalytic performance (Bordeaux et al., 2011, Munro et al., 2007, Bakkes et al., 2017). CYPs are an enzyme superfamily and therefore there are many to choose from. CYPs are selected according to their unique characteristics or benefits that suit the application of interest. In the present study, CYP153A6 and CYP153A13 were chosen for reasons described below.

The CYP153A6 multicomponent system consists of three separate components- the CYP enzyme and its electron transfer partners ferredoxin (Fdx) and ferredoxin reductase (FdR) (Funhoff et al., 2006, Bakkes et al., 2017) (Figure 1.1). The spatial arrangement of multicomponent systems increases the chance that electrons will be uncoupled or transferred slowly. Electrons are more likely to contribute to a side-reaction and not the alkane oxidation reaction of interest, creating oxidants that deactivate the active site (Bakkes et al., 2017). Alternatively, self-sufficient fusion CYPs are expressed as a single component and comprise electron transfer proteins fused to the heme domain of the CYP (Bordeaux et al., 2011, Munro et al., 2007) (Figure 1.1). This arrangement allows for improved coupling efficiency and self-sufficiency. Biotechnological tools and protein engineering allows for the fusion of the P450 enzyme to its natural or foreign reductase partners to overcome slow or decoupled electron redox transfer (Nodate et al., 2006).

CYP102A1 is a well-known soluble fusion enzyme with a higher activity than CYP153s, but with lower selectivity. The respective traits of high activity of soluble fusion CYPs and high selectivity of CYP153s were utilised by constructing a fusion CYP known as the CYP153A13 (Figure 1.1B). Kubota et al. (2005) constructed the artificial fusion enzyme CYP153A13 by fusing CYP153a from *Alcanivorax borkumensis* SK2 to RhFred- a flavin mononucleotide (FMN)/Fe_2S_2-containing reductase domain from *Rhodococcus* sp. NCIMB 9784. The fusion CYP153A13 shows high selectivity for the terminal carbon of linear alkanes and produces primary alcohols, fulfilling a niche within the field of biocatalysis.

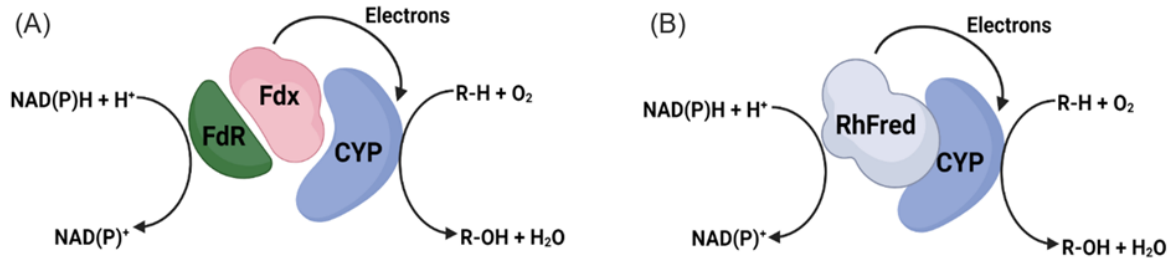


Figure 1.1 Schematic diagram of the structural organisation of cytochrome systems. (A) Multicomponent enzyme consisting of the CYP enzyme and two electron transfer partners FdR and Fdx. (B) Self-sufficient fusion enzyme consisting of CYP heme domain fused to an electron transfer protein such as RhFred.

CYP153A13 previously showed promising binding affinities measured by the affinity binding constant K_D to medium to longer chain alkanes: $K_D = 1\ \mu M$ for *n*-dodecane, $K_D = 16\ \mu M$ for *n*-decane and $K_D = 29\ \mu M$ for *n*-octane (Bordeaux et al., 2011). CYP153A13, however, has shown to suffer from poor expression, potential low solubility and functional expression compared to CYP102A1 (Bordeaux et al., 2014). Despite the several advantages of CYP153A13, the CYP153A6 multicomponent system has shown superior and higher expression levels of functional and soluble protein (Funhoff et al., 2007, Bordeaux et al., 2014). The multicomponent CYP153A6 system however comprises of separate redox partner proteins, causing slower electron coupling (Munro et al., 2007). CYP153A6 and CYP153A13 were selected because both of these enzymes have shown promising activity, high hydroxylation selectivity and affinity for the substrate of interest in this study- C_8, C_9 and C_{10} alkanes (Funhoff et al., 2006, Kochius et al., 2018, Bordeaux et al., 2011). Furthermore, the selection of these two enzymes enabled the comparison of a multicomponent enzyme to a fusion enzyme.

1.4 Whole cell biocatalysis for alkane hydroxylation

Biocatalytic approaches can employ either the crude or pure enzyme preparations or whole cells containing the enzymes. Whole cells have a higher operational stability because they provide protection of the enzymes against harsh reaction conditions and oxidative stress. Other advantages of whole cell biocatalysis is that the cell regenerates expensive NAD(P)H needed for the alkane activation reaction, rendering whole cells as 'self-sufficient'. Whole cells are cheaper than isolated enzymes because there is no need for expensive enzyme purification steps (Wachtmeister and Rother, 2016, Schrewe et al., 2013).

Whole cell biocatalysts are sensitive to harsh reaction conditions, limiting the range of operation parameters. Whole cell biocatalysts are stable and reside in aqueous media, while linear alkanes are hydrophobic and have poor water solubility with limited access to the biocatalyst (Schmid et al., 1998). The 2-Liquid-Phase-System (2LPS) consisting of an aqueous liquid media and organic liquid phase is commonly employed to curtail these limitations. The organic phase consists of an organic carrier solvent that acts as a substrate reservoir and product sink. This allows placement of high substrate loadings and regulation of toxic substrate/product concentrations in the aqueous phase (Bühler et al., 2003). Low water solubility of medium to longer chain alcohols causes phase separation and consequently allows for easier product extraction, reduced toxicity to microorganisms and lower recovery costs (Lennen and Pfleger, 2013).

Poor solubility of medium to longer chain linear alkanes prevents catalytic activity to proceed at maximum enzyme functionality and activity because the concentration of substrate in the aqueous phase is limited (Eckstein et al., 2006a). Ideally, thermodynamic equilibrium should favour the delivery of high concentrations of substrate to the aqueous media and efficient extraction of product into the organic solvent. The reaction equilibrium and phase equilibrium of the 2LPS can be shifted by altering physicochemical parameters such as the nature of the solvent, initial substrate concentration and phase ratio (Kim et al., 2007).

The kinetics of an enzyme-catalysed reaction are typically obtained by the measurement of discrete initial reaction velocities (v_0) from a series of experimental runs with different substrate concentrations. Substrate concentration versus v_0 data points are plotted to derive the hyperbolic Michaelis-Menten (MM) model. Consequently, kinetic parameters are obtained namely the maximum velocity (v_{max}) and the MM constant K_m which is the substrate concentration at which half the v_{max} is reached.

The K_m is used as an indicator of the enzyme-substrate (ES) binding affinity (Palmer and Bonner, 2007, Robinson, 2015). In the case of whole cell 2LPSs, kinetic parameters such as the measured K_m and v_0 are dependent not only on the turnover of the reaction and ES interaction, but also substrate-water molecular interaction, solvent-substrate interaction, transmembrane transport, interphase mass transfer and partitioning of substrate/product. These experimentally measured or 'apparent' kinetic parameters include interactions between both ES as well as solvent-substrate, assumes that the enzyme is exposed to excess substrate and do not reveal the intrinsic or true ES affinities and kinetics (Ferrario et al., 2018).

The present study aimed to evaluate the kinetics of the CYP153 whole cell catalysed oxidation of medium chain alkanes in terms of substrate solubilities. Substrates of different solubility were compared using thermodynamic activity-based kinetic analyses to separate indirect effects of solvent-substrate from true ES affinity. Whole cell biocatalytic primary hydroxylation performance of the well-studied multicomponent CYP153A6 enzyme system was compared to the fusion CYP153A13 enzyme on a range of linear medium-chain *n*-alkanes. Fusion enzymes are known to have superior coupling efficiencies compared to multicomponent enzymes, but also show difficulties in expression.

Biocatalytic processes could be significantly improved by identifying which of the several inter-dependent factors limit productivity. The present study hypothesises that, for medium chain alkane biohydroxylation, substrate availability, -delivery and -access to the biocatalyst is one of the most significant constraints on biocatalytic activity. Limitation of substrate access to the biocatalyst arises from parameters such as mass transfer, substrate solubility in the aqueous phase and substrate uptake into the whole cell membrane. Although this may be one of the most significant constraints, other factors such as enzyme expression and electron transfer rates also play a role in overall biohydroxylation performance. Moreover, it was hypothesised that the decreased aqueous solubility of alkanes of longer chain length results in lower reaction rates. By using thermodynamic activity-based kinetic analyses, it is possible to decouple solvent-substrate from ES interactions to reveal intrinsic enzyme kinetics on a particular substrate and to reveal the preferential substrate for CYP153A6 and CYP153A13, respectively (Ferrario et al., 2018).

1.5 Research objectives and aims

The 2LPS primary hydroxylation reaction kinetics of whole cell biocatalysts expressing CYP153A6 and CYP153A13 on *n*-octane, *n*-nonane and *n*-decane were investigated. The overall biocatalytic performance and reaction kinetics were evaluated in terms of constraining factors such as substrate supply to the whole cell biocatalyst residing in the aqueous phase and functional protein expression

levels. Experimentally measured substrate concentration-based apparent reaction kinetics were investigated. To determine the most-well suited ES combinations, thermodynamic-related solvation effects were accounted for. The aim was to gain insight into physicochemical and enzyme attributes of 2LPS biocatalytic hydroxylation processes. This enabled the determination of the ES reaction rates that were limited by substrate availability, and if substrate was not limiting, other factors such as enzyme-substrate affinity and enzyme concentration were considered.

The performance of the multicomponent CYP153A6 system was compared to the fusion CYP153A13 enzyme to explore potential benefits that fusion enzymes may offer due to enhanced electron coupling. This was done by characterising the expression of both enzyme systems. The objective was to identify the limitations of experimental reaction kinetics with respect to enzyme expression, substrate solubility and availability to the whole cell biocatalyst and ES affinity. The analyses of CYP153 whole cell kinetics on the hydroxylation of alkanes is imperative for understanding scalability of 2LPSs and accurate selection of the most suited biocatalysts. This study also unified concepts from both molecular biology and chemical engineering science. This multidisciplinary approach is critical for industrial development of biocatalysis.

1.6 Dissertation structure

This dissertation is introduced with a detailed background review of existing knowledge within the field of biocatalytic alkane hydroxylation. in the literature review Chapter 2.1 covers conventional routes to alkane activation and is followed by Chapter 2.2 in which various alkane hydroxylase enzymes and their mechanisms are introduced. Chapter 2.3 details the various biotechnological tools applied in the field of biocatalysis. Chapter 2.4 discusses the characterisation of CYPs and their stability, expression, and activity. Chapter 2.5 describes process and operating conditions central to alkane biohydroxylation, physicochemical, thermodynamic, and mass transfer considerations of 2LPSs and the nature of alkane substrates. The literature review is concluded with project hypotheses, objectives, and key questions. Chapter 3 presents the methods that were used to answer project key questions. The results and findings are illustrated, interpreted, and discussed in Chapter 4. Chapter 5 follows with overall project conclusions related to the stated objectives, as well as findings of the project and its recommendations for future work.

2 Literature Review

Alkanes are well-known for their inertness due to their completely saturated nature and due to their unactivated C-H bonds that possess 400 $kJ \cdot mol^{-1}$, but do still react with oxygen present in air under certain conditions (Shilov and Shul'pin, 1997). Combustion of alkanes can provide a source of energy as well as the selective production of thermodynamically favourable products carbon dioxide and water. Alkanes are far more valuable however when used as precursor molecule feedstock for the synthesis of fine chemicals (Labinger and Bercaw, 2002). The next section will cover the different routes to the functionalisation of alkanes. There are three major routes to activate alkanes: heterogenous catalysis, organometallic activation, and biological/biomimetic catalysis.

2.1 Conventional routes to alkane activation

Autooxidation or better known as “cracking” of alkanes can occur at high temperatures (>1000˚C), but generally these reactions produce undesirable products and suffer from poor selectivity. Some of the first alkane activation reactions were catalysed by transition metal complexes such as Fenton’s reagent, an example of homogenous catalysed alkane activation (Fenton, 1894). In the 1970s, developments in this field focused on platinum(II) salt catalysed C-H activation (Labinger, 2016, Crabtree and Lei, 2017).

Organometallic based catalysis suffers from the degradation of metal-complexes under the required oxidising reaction conditions. The Shilov system solved this problem by adding Pt(IV) ionic solutions to water to oxidise alkanes to alcohols at 120˚C (Shilov and Shul'pin, 1997). From the 1980s onwards, transition-organometal (especially Pt) based catalysis of C-H activation was comprehensively studied. Some other examples of inorganic and organometallic catalysts include shape-selective metalloporphyrins, Lewis acid-assisted metal compounds, immobilized mononuclear iron carboxylates and shell hydroformylation technology. Other routes to alkane activation include heterogenous catalysis and biomimetic catalysis (Olaofe et al., 2013).

Conventional routes such as organometallic catalysis to alkane activation offer higher conversion rates and are applied at large-scale, they suffer from poor selectivity, over-oxidised product, and unwanted side product formation. Most chemical catalysts lose activity due to harsh oxidants needed for the reaction. These factors increase downstream processing costs and process efficiency.

2.2 Biocatalysis

Living organisms all contain metabolisms necessary for homeostasis and cellular functioning. Metabolic pathways consist of networks of biochemical reactions catalysed by proteins called enzymes. It was discovered that enzymes found in nature can be used in the application of biocatalysis for complex reactions and for the artificial synthesis of valuable chemicals. Biocatalysis is emerging as a practical alternative approach of alkane activation to enable highly selective alkane activation under milder process and operation conditions, with reduced side-product formation. The high product selectivity and wide substrate compatibility make biocatalysis an attractive approach for alkane oxidation.

Compared to conventional chemical cracking, biocatalytic fine chemical production has the potential to reduce toxic solvent usage, production costs, downstream processing costs, energy consumption and environmental burden. Indeed, biocatalysis shows promising potential for transition of fine chemical production industry towards sustainable and environmentally conscious practices. The bottlenecks of enzyme-based catalysis, however, are vast, inter-dependent and complex, especially since the function

of enzymes evolved for organism survival and not for mass large-scale production of fine chemicals (Ayala and Torres, 2004).

2.2.1 Alkane hydroxylases

Alkane hydroxylases are a family of enzymes recognised for their ability to catalyse the stereo- and regioselective oxygenation of highly reduced hydrocarbons. Alkane hydroxylases are found in a wide range of organisms but were first discovered and isolated from alkane degrading microbial species found in oil-impacted environments (Van Beilen and Funhoff, 2007). Alkane hydroxylases play a major role in the β-oxidation pathway, which serves to generate energy for the cell from fatty acids (Nie et al., 2014a). These enzymes oxidise alkanes to primary alcohols, then further oxidise primary alcohols to fatty acids, which then enter the β-oxidation pathway (Nie et al., 2014a, Bordeaux et al., 2012). The discovery of alkane hydroxylases prompted their use to achieve one of the most challenging reactions known to organic chemistry - alkane activation to produce more reactive, oxygenated and higher-value primary alcohols (Van Beilen and Funhoff, 2007, Koch et al., 2009b, Ron and Rosenberg, 2002).

There are a few defining characteristics shared among the members of this enzyme family. Firstly, they all contain a copper or iron-based active site which catalyses the oxidation of alkanes under physiological conditions. This oxidation reaction occurs by concomitant reduction of molecular dioxygen, with NAD(P)H as an electron supplier (Equation 2.1).

$$R\text{-}H + O_2 + NAD(P)H + H^+ \xrightarrow{\textit{Alkane hydroxylase}} R\text{-}OH + NAD(P)^+ + H_2O \qquad \textit{Equation 2.1}$$

Categorisation of alkane hydroxylases are based on alkane degradation characteristics and preferred substrate chain length. These include methane mono-oxygenases, integral membrane non-heme diiron alkane ω-hydroxylases (AlkB) and cytochrome P450 enzymes (CYP) (van Beilen and Funhoff, 2005, Bordeaux et al., 2012, Ji et al., 2013). Table 2.1 summarises the different alkane hydroxylases that have been applied in various alkane activation studies.

Table 2.1. Different alkane hydroxylase groups, their wild-type species of origin, cellular locations, preferred substrates and previously reported catalytic rates

Alkane hydroxylase	Wild-type species	Cellular location	Compatible substrates	Catalytic rate on alkanes (min^{-1})	Reference
Soluble methane monooxygenase (sMMO)	*Methylisinus trichosproium, Methylococcus capsulatus*	Soluble Cytoplasmic	C_1-C_8	200-250	(van Beilen and Funhoff, 2005, Urlacher et al., 2004)
Particulate methane monooxygenase (pMMO)	Methanotrophic bacteria	Membrane bound	C_1-C_5	200-325	(van Beilen and Funhoff, 2005, Lieberman and Rosenzweig, 2004)
Integral membrane non-heme alkane hydroxylases (AlkB-related)	*Acinetobacter*, *Alcanivorax*, *Mycobacterium, Pseudomonas, Rhodococcus*	Membrane bound	C_5-C_{10}	210	(Táncsics et al., 2015, Rojo, 2005)
Eukaryotic P450 (CYP52)	*Candidia maltose, Candida tropicalis, Yarrowia lipolytica*	Membrane bound	C_{10}-C_{16}	27-40	(Scheller et al., 1996)
Prokaryotic P450 (CYP153)	*Sphingomonos, Mycobacterium, Acinetobacter, Alcanivorax borkumensis SK2*	Soluble Cytoplasmic	C_5-C_{12}	0.1-70	(Funhoff and Van Beilen, 2007, Fujita et al., 2009, Kubota et al., 2005)
Prokaryotic P450 (CYP102A1)	*Bacillus megaterium*	Soluble Cytoplasmic	C_2-C_{10}	860-3020	(Eiben et al., 2006)
LadA	*Geobacillus thermodenitrificans*		C_{15}-C_{36}	4.4	(Van Beilen et al., 2003b, Scheps et al., 2011)

2.2.1.1 Methane mono-oxygenases

Methanotrophic microbial species contain methane mono-oxygenases that allow oxidation and utilisation of methane as an energy source (Lieberman and Rosenzweig, 2004). Membrane-bound particulate methane monooxygenases (pMMOs), iron-copper enzymes, are expressed when there is high copper to biomass ratio present in the environment. Some methanotrophic species evolved to live in environments of low copper levels by expressing di-iron based soluble MMOs (sMMOs) (Ji et al., 2013, van Beilen and Funhoff, 2005, Ayala and Torres, 2004). sMMOs have a wider substrate range (including methane to octane, alkenes, aromatic, halogenated hydrocarbons) than pMMOs, which are limited to short chain alkanes and alkenes (Ayala and Torres, 2004, Bordeaux et al., 2012). The application of MMOs in biocatalytic alkane activation however has been limited due to their multicomponent nature, difficulties in recombinant expression and characterisation, product inhibition susceptibility, instability and low selectivity (Bordeaux et al., 2012, Koch et al., 2009b).

2.2.1.2 Integral membrane non-heme alkane hydroxylases

One of the first and most well-studied alkane hydroxylase systems was discovered in the hexane-degrading *Pseudomonas putida* GPo1. This system consists of three components: an integral-membrane non-heme diiron monooxygenase (AlkB) and its two electron transfer proteins, rubredoxin (AlkG) and rubredoxin reductase (AlkT) (Shanklin et al., 1997, Meinhold et al., 2006). AlkL is an outer membrane transport protein from *P. putida* shown to facilitate uptake of alkanes into the cell (Julsing et al., 2012)

Additional AlkB-related genes were later discovered in a wide range of β- and γ- *Proteobacteria* and *Acinetobacteria* species (Funhoff and Van Beilen, 2007, Nie et al., 2014a). Species expressing Alk genes can use medium-length alkanes (C_6-C_{10}) as their sole energy and carbon source (Ayala and Torres, 2004). Members of the AlkB-related enzyme family contain a diiron cluster and act by using reducing equivalents from NAD(P)H or NADH to oxidise alkane chain lengths from C_5 up to C_{16} at the terminal carbon (Koch et al., 2009b). These enzymes attracted attention for their wide substrate range, including medium to longer chain length alkanes, as well as their reasonably high turnover frequencies (200 $\mu mol_{1-octanol} \cdot \mu mol_{biocatalyst}^{-1} \cdot min^{-1}$) (Shanklin et al., 1997). The Alk family of enzymes are membrane-bound and multicomponent enzyme, which presents difficulties for in vitro studies, recombinant techniques, and biocatalytic application (Ayala and Torres, 2004, Bordeaux et al., 2012).

2.2.1.3 Flavin-containing alkane mono-oxygenases

Flavin-containing alkane mono-oxygenases (LadA) are capable of oxidising longer chain alkanes ($\geq C_{16}$) (Scheps et al., 2011). The LadA enzyme was isolated from a bacterial strain *G. thermodenitrificans* NG80-2 that metabolises long chain alkanes (C_{15}-C_{36}). LadA requires partner proteins and cofactors FMN or NAD(P)H for alkane oxidation. A distinct feature of these enzymes is their metallic-free active centre and alkane oxidation mechanism (Bordeaux et al., 2012).

2.2.1.4 Cytochrome P450s

CYPs are hemoproteins found in a wide range of organisms. More than 4000 CYP enzymes have been discovered across a broad range of species ranging from eukaryotic species such as plants and mammals to more simple prokaryotic species such as bacteria (van Beilen and Funhoff, 2005). These enzymes play a role in the metabolic oxidation of a wide range of aliphatic compounds such as xenobiotics, steroids, fatty acids and prostaglandins (Bordeaux et al., 2012, Urlacher et al., 2004). CYPs are the major biochemical catalysts for the β-oxidation pathway to metabolise fatty acids. Several naturally occurring CYPs capable of hydroxylating alkanes have been discovered and studied.

Comprehensive studies of CYPs have led to a deeper insight into their functioning, mechanistic catalytic action, and gene regulation, amongst many more.

Only a few CYPs that hydroxylate *n*-alkanes have been identified with the potential for industrial use (Ji et al., 2013). Characteristics of industrially relevant CYPs include: soluble, amenable to genetic manipulation, self-sufficient, high catalytic rates, turnover numbers and selectivity, stable, mild operative conditions, and minimal downstream processing (Funhoff and Van Beilen, 2007). For example, CYP102A1 shows catalytic rates on C_2 to C_{10} alkanes as high as 3020 min^{-1}. This was attributed to increased coupling efficiency due to the closer spatial arrangement the redox partners to the heme group of a fusion CYP, compared to multicomponent CYPs (Ciaramella et al., 2017).

Hydroxylation reactions catalysed by CYPs occur via the heme active site. The reaction requires redox partner protein/s, cofactors NADH or NADPH and molecular oxygen (Urlacher and Girhard, 2012). Electrons required for the hydroxylation reaction are transferred from NADPH or NADH via redox partner proteins to the P450 heme active site. The active site then reduces one oxygen atom of molecular oxygen into water and selectively introduces the other oxygen atom into the hydrocarbon substrate at the primary carbon position. The overall efficiency at which electrons are transferred to the alkane substrate from the electron donor cofactor is termed 'coupling efficiency'. Coupling efficiency of a biocatalytic alkane oxidation reaction indicates the efficiency of donated electrons to the reaction of interest (i.e., primary alkane oxidation). Uncoupling occurs when electrons are transferred to molecules such as water to produce hydrogen peroxide, an unwanted by-product that is known to deactivate the heme prosthetic group active site of CYPs (Bordeaux et al., 2011).

Self-sufficient fusion CYPs consist of an alkane hydroxylase unit fused to redox protein partners and are expressed as a single polypeptide chain (Ciaramella et al., 2017). According to Ciaramella et al. (2017), the fusion CYP102A1 has a strikingly high activity of 17 000 s^{-1}. Reports have shown that fusion proteins have high turnover numbers (up to 15 000 s^{-1}), which makes them attractive as industrial biocatalysts. Fusion proteins simplify alkane oxidation reactions because they require only NAD(P)H and substrate (Ciaramella et al., 2017). Multicomponent CYP systems such as the CYP153A6 require the expression of additional and separate redox protein partners and efficient and coupled electron transfer rates along these components (Bordeaux et al., 2011, Eiben et al., 2006).

The first naturally occurring self-sufficient and soluble P450 discovered was CYP102A1 from *Bacillus megaterium* (P450 BM3) which hydroxylates C_{12}-C_{18} fatty acids and has become the model fusion P450 enzyme (Ciaramella et al., 2017, Ayala and Torres, 2004). The high activity, high expression, and single-component nature of CYP102A1 are desirable characteristics of an industrial biocatalyst for alkane hydroxylation. The CYP102A1 family however has shown low stereo- and regio-selectivity for terminal alkane oxidation compared to the CYP153 family (Funhoff and Van Beilen, 2007). Another promising fusion enzyme that was discovered is CYP116B2 (also known as P450RhF) isolated from *Rhodococcus* sp. NCIMB 9784. CYP116B2 is expressed as a single polypeptide that consists of three domains: FMN-binding domain, NAD(P)H-binding domain and a [2Fe-2S] ferredoxin-like centre (Ciaramella et al., 2017, Nodate et al., 2006).

Soluble CYPs can be recombinantly expressed, showing potential for industrial application (Ortiz de Montellano, 2009). Fungal, mammalian and plant P450s are membrane bound, more difficult to handle, low in activity and are not easily expressed recombinantly or genetically manipulated. Yeast species such as *Candida tropicalis* express microsomal CYP52s (Bordeaux et al., 2012). Similarly to MMOs and AlkBs however, microsomal CYPs are limited in biocatalytic application in engineered systems because they are membrane-bound, multicomponent, exhibit low activity and show difficulties in

recombinant expression (van Beilen and Funhoff, 2005, Bordeaux et al., 2012, Ayala and Torres, 2004, Ji et al., 2013). This limits their use to wild-type systems.

On other other hand, many bacterial CYPs are soluble and are easily manipulated, making these an attractive choice of enzyme for biocatalytic alkane activation for industrial application. Most CYPs from prokaryotic species are soluble, which offers significant advantages for biocatalytic application but most of these have shown no activity towards *n*-alkanes (Funhoff and Van Beilen, 2007, Maier et al., 2001). In 2001, the first soluble CYP that displayed hydroxylation activity towards linear *n*-alkanes was purified from the alkane-degrading bacterial *Acinetobacter* species and characterised (CYP153A1) (Maier et al., 2001). Gene sequence alignment studies revealed that CYP153A1 had low similarity to other known CYP families.

The establishment of the CYP153A1 sequence initiated the search for other members of the CYP153 family (Funhoff et al., 2006, van Beilen et al., 2006). For example, the discovery of the CYP153A6 in the *Mycobacterium* sp. HX-1500 was noteworthy for its cytosolic location and solubility and ≥95% hydroxylation regioselectivity of the terminal carbon of alkanes (*n*-pentane to *n*-dodecane) (Ji et al., 2013, Bordeaux et al., 2012, Funhoff et al., 2006). Funhoff and Van Beilen et al. (2007) performed substrate binding studies which entailed determining the binding/dissociation constant K_D. The K_D constant indicates the affinity of an enzyme for a substrate and how well the shape of the active site of the enzyme complements the substrate, discussed in detail in Chapter 2.5.3.2. Funhoff and Van Beilen et al. (2007) obtained promising K_D constants for CYP153A6 on medium chain linear alkanes, such as a K_D of 20 nM for *n*-nonane.

CYP153s have become pivotal in the field of biocatalytic hydroxylation of medium to longer chain-length *n*-alkanes under mild operation conditions (van Beilen and Funhoff, 2005, Bordeaux et al., 2012). The CYP153s can hydroxylation at the terminal carbon with high selectivity, unlike AlkB and other P450 families (Olaofe et al., 2013). CYP153s show higher selectivity for terminal carbon hydroxylation than other CYP families including CYP102A1. Studies on CYP153s have focused on characterisation, purification, expanding substrate specificity and improving activity and ease of handling (Bordeaux et al., 2012).

Widespread industrial and scalable application of P450s is limited due to poor substrate specificity, low activity and stability (Urlacher and Girhard, 2012). Some bottlenecks of biocatalysis include requirement of additional redox protein partners, expensive cofactor requirements, insolubility, low activity, difficulties in protein expression and improper protein folding. Wild-type enzymes provide a basic platform to engineer biocatalysts suitable for industry. Protein engineering and biotechnology allows for the design and modification to improve on and acquire enzymes that exhibit industrially relevant and appealing characteristics (Bordeaux et al., 2012).

2.3 Biotechnological tools for biocatalysis

Naturally found CYPs often either display high activity but low specificity or vice versa. Therefore natural CYPs often do not exhibit multiple desirable traits in a single system (Eiben et al., 2006). The improvement on naturally occurring catalysts through their engineering to achieve custom-made biocatalysts is possible through technological advancements such as DNA sequencing, DNA plasmid construction, recombinant expression, protein engineering and metabolic engineering. Recombinant protein expression, as well as protein and metabolic engineering techniques can be used to create biocatalysts with novel and improved properties such as higher activity, regioselectivity, stability,

substrate specificity, solubility, self-sufficiency and electron coupling efficiencies (Eiben et al., 2006, Bordeaux et al., 2011).

Enzymes can be modified by either site-directed mutagenesis or directed evolution (O'Reilly et al., 2011). Some examples include the creation of artificial fusion enzymes inspired by natural fusion enzyme design, combining subunits with desirable traits from different enzymes, modification of substrate binding sites to expand substrate specificities and thereby improve enzyme activity (Eiben et al., 2006, Bordeaux et al., 2011, Yang et al., 2014, Ciaramella et al., 2017). Substrate specificity can be altered by replacing certain amino acid residues present in the active site of an enzyme. For example, the P450cam system (also known as CYP101) from *Pseudomonas putida* ATCC17453 naturally oxidises camphor. The camphor-binding pocket was modified by rational design and amino acid substitution with larger hydrophobic residues, causing a reduction in the substrate-binding pocket volume, which improved its compatibility toward *n*-alkane substrates reaching a turnover number of 1000 min^{-1} on *n*-butane (Ji et al., 2013, van Beilen and Funhoff, 2005, Bell et al., 2003).

The discovery and comprehensive characterisation of the naturally occurring CYP102A1 soluble fusion system prompted the creation of artificial fusion proteins (Munro et al., 2007). CYP102A1 naturally hydroxylates middle to long chain saturated and unsaturated branched fatty acids at the subterminal carbon (Eiben et al., 2006). Arnold and co-workers used protein engineering techniques to obtain a CYP102A1 mutant capable of hydroxylating short linear alkanes. Despite several rounds of directed evolution and studies, hydroxylation mainly occurred at sub-terminal positions (Meinhold et al., 2006). CYP102A1 is more active than CYP153 enzymes, but less regio-selective.

Kubota et al. (2005) cloned various CYP153 hydroxylase genes fused to the reductase domain of CYP116B2 (P450RhF) from *Rhodococcus* sp. NCIMB 9784 (known as RhFred) into *Escherichia coli*. In this study, a novel artificially engineered self-sufficient fusion P450 was made by fusing the CYP153A13a heme domain (P450balk) from *Alcanivorax borkumensis* SK2 to an FMN(Flavin mononucleotide)/Fe_2S_2-containing redox domain of CYP116B2 (P450RhF) from *Rhodococcus* sp. NCIMB 9784 (Kubota et al., 2005, Nodate et al., 2006). This CYP153A13 artificial fusion system was expressed, purified, and characterised *in vitro*, demonstrating the remarkable regio- and chemoselectivity on medium alkanes to produce primary alcohol products under mild conditions (Bordeaux et al., 2011). A subsequent study demonstrated that CYP153A13 suffers from operational instability and difficulties in recombinant expression (Bordeaux et al., 2014).

2.4 Biocatalyst characterisation

2.4.1 Enzyme production, activity, and stability

Biotechnological efforts to improve intrinsic enzyme properties or expression have proven valuable to the field of biocatalytic alkane activation, but processes using even customised engineered CYP153s are still constrained by poor productivity, short reaction times, low stabilities and low activities (Duetz et al., 2001). This directs attention toward constraining parameters that are beyond intrinsic enzyme properties. Several factors affect protein expression and functioning which, in turn, affect biocatalytic activity and stability. These factors range from molecular variables that affect enzyme functioning to macro-variables such as process conditions that affect biocatalyst functioning. The major parameters used to report the overall biotransformation performance of biocatalytic systems are summarised in Table 2.2. Biocatalytic characterisation approaches to report key performance parameters are different for cell-free and whole cell systems. For cell-free processes, turnover frequency (TOF) is the quantity (mole) of product yielded per quantity (mole) of enzyme used per time unit. In contrast the TOF for a

whole cell processes is mole of product yielded per gram of dry cell weight per time unit. Sometimes, whole cell processes can also be measured relative to the quantity of enzyme if methods are available for accurate measurement. Matters are further complicated by inconsistent reporting (terms, units) of biocatalytic data across different contributing laboratories and different disciplinary backgrounds, resulting in a lack of comparability and reproducibility (Gardossi et al., 2010).

Table 2.2. Key performance parameters of cell free and whole cell biocatalytic systems (Gardossi et al., 2010, Schrewe et al., 2013)

	Cell-free		Whole cell	
Measurement	Variable	Unit	Variable	Unit
Rate	Volumetric productivity rate, productivity, or space time yield	$mol_{product} \cdot L_{WV}^{-1} \cdot min^{-1}$	Volumetric rate, productivity, or space time yield	$mol_{product} \cdot L_{WV}^{-1} \cdot min^{-1}$
	Specific Activity or Turnover Frequency (TOF)	$U \cdot g_{enzyme^{-1}}$ or $U \cdot \mu mol\ P450^{-1}$	Specific Activity or Turnover frequency	$U\ g_{DCW^{-1}}$ or $U \cdot \mu mol\ P450^{-1}$
	Maximum Specific Biocatalytic Activity			$mol_{product} \cdot g_{DCW}^{-1} \cdot min^{-1}$
Yield	Specific Product Concentration or Turnover Number (TON) or Total turnover number	$mol_{product} \cdot mol_{enzyme^{-1}}$	Product yield on catalyst, metabolic flux, or TON	$g_{product} \cdot g_{DCW^{-1}}$ $mol_{product} \cdot g_{DCW^{-1}}$
Yield	Product yield on substrate consumed ($Y_{P/S}$)	$mol_{product} mol_{substrate^{-1}}$	Product yield on substrate	$mol_{product} \cdot mol_{substrate^{-1}}$
Biocatalyst Efficiency	Biocatalyst Efficiency	$g_{product} \cdot g_{P450}{}^{-1}$		$g_{product} \cdot g_{DCW}{}^{-1}$
Specific Intracellular P450 Concentration	P450 concentration	n.a	Whole cell P450 concentration	$\mu mol_{P450} \cdot g_{DCW}{}^{-1}$
Affinity	K_D	$mol_{substrate} L^{-1}$	K_D	$mol_{substrate} \cdot L^{-1}$

$U = \mu mol_{product} \cdot min^{-1}; DCW = dry\ cell\ weigh;\ L_{WV} = Working\ volume\ available\ for\ reaction$

Among the reasons for preferring whole cell systems versus cell-free systems are increased enzyme stability as well as co-expression of redox partners. The choice of host cells and recombinant expression systems have been shown to significantly influence overall bioprocess productivity (Lundemo and Woodley, 2015). The use of wild-type species containing natural alkane hydroxylase systems is limited in industrial application for several reasons, including the practical difficulties of culturing most alkane-utilising species and that alkane hydroxylases are expressed at low levels in wild-type hosts which limits bioconversion performance.

Bordeaux et al. (2011) concluded that CYP153A13 suffers from poor expression resulting in low concentrations and this results in variability of productivity of CYP153A13 recombinant cultures and sometimes even observing unproductive cultures (Bordeaux et al., 2011, Bordeaux et al., 2014). Another study also stated that fusion proteins often suffer from poor expression (LaVallie and McCoy, 1995). Bordeaux et al. (2014) used High Cell Density Cultures (HCDC) to compensate for low CYP153A13 expression levels by increasing the cellular density per unit volume of BRM and thereby increase the mole amount of CYP153A13 present per unit volume of BRM. These authors argued that active CYP153A13 concentration is low per cell and therefore by increasing cell concentration, this would increase the concentration of active and functional CYP153A13 in the total culture volume i.e. total volumetric concentration of CYP153A13.

Protein engineering allows for expression of desired protein in foreign hosts such as *E. coli* to allow for easier cultivation and higher protein expression levels (Gudiminchi et al., 2012). Hosts such as *E. coli* do not possess intrinsic genes to express CYP but *E.coli* is the most extensively applied expression host in CYP production (Donovan et al., 1996, Lundemo and Woodley, 2015). Several expression vectors have been developed for stable recombinant protein production in *E. coli* (Fujita et al., 2009, O'Reilly et al., 2011). The vast knowledge of genetic manipulation and optimal recombinant expression in *E. coli* makes it an attractive choice for whole cell biocatalysis (Makrides, 1996, Donovan et al., 1996). Further, *E. coli* grows on inexpensive media, to high-densities and is commonly used for research.

While many techniques have been developed for optimal and high expression of recombinant protein in *E. coli* to support efficient biocatalytic activity, over-expression of enzyme does not necessarily result in higher biocatalyst activity (Gudiminchi et al., 2012). Nevertheless, *E. coli* is the preferred host compared to hosts such as yeast which is more complex or the *Pseudomonas* species which is not Generally Regarded As Safe or GRAS (Kampers et al., 2019).

To optimise protein expression, the promoter of the operon consisting of the P450 component, and its redox partners should be tightly controlled and simple and cost effective to induce. Metabolic burden is reduced if the times and levels of protein induction are controllable (Olaofe et al., 2013, Makrides, 1996). Regulatable promoters express the protein of interest in response to a specific controllable environmental change, such as an added chemical constituent (Donovan et al., 1996). The *lac* promoter, for example, can be induced in the presence of isopropyl-β-D-thiogalactopyranoside (IPTG) or lactose. The IPTG system confers many advantages over using lactose because it is not involved in basal cell metabolism (i.e., it is a gratuitous inducer). For the purposes of a small-scale laboratory process requiring as little metabolic interference as possible, the use of the commonly used, developed and powerful IPTG induction system is advantageous (Donovan et al., 1996, Makrides, 1996).

Once expressed, the enzyme acquires stability via proper protein folding into a ‘quaternary structure’ to display activity (Palmer and Bonner, 2007). Regulation of the rate of protein synthesis by lowering postinduction temperature for example, ensures proper folding of protein and in turn prevents the formation of inclusion bodies (Makrides, 1996, Lee, 1996). The quaternary structure of a protein is stabilised by disulphide bridge bonds between cysteine residues, electrostatic interactions or van der Waals bonds. Non-polar or uncharged amino acid side chains are often packed to form a hydrophobic core, while polar and charged amino acid side chains form a hydrophilic outside of the enzyme. Biocatalyst stability majorly relies on the maintenance of the quaternary structure of the enzyme in the reaction conditions for the duration of the reaction (Bommarius and Paye, 2013). Overly rapid expression interferes with the protein folding process and therefore can result in inclusion bodies.

2.4.2 Mechanism of CYP153 biocatalytic alkane activation

Hydroxylase activity is cofactor dependent, and the oxidation of alkane substrate requires consumption of NAD(P)H to provide electrons to the heme active site (Meyer et al., 2006). The heme active site of CYPs is unable to accept electrons directly from NAD(P)H and therefore requires the assistance of electron transfer proteins. Therefore, CYP153 activity relies on not only its optimal expression and stability, but also the coupling efficiency of the electron transfer system to form product (Scheps et al., 2011, Bordeaux et al., 2011).

Transfer of electrons from NAD(P)H to the active site of the multicomponent CYP153A6 system occurs via electron transfer proteins ferrodoxin reductase (FdR) and ferrodoxin (Fdx). Upon correctly orientated substrate binding to the heme active site of the CYP, the substrate-bound complex increases its ability to accept electrons from the redox partners. The reduced active site selectively introduces one oxygen atom into the hydrocarbon substrate at the primary carbon position and the other oxygen atom is reduced to water.

Biocatalytic efficiency can be hampered by slow or uncoupled electron transfer (Bordeaux et al., 2011). The redox partners that assist the electron transfer, the rate at which this transfer occurs as well as the correct utilisation of the electrons from the cofactor to the substrate of interest (known as the coupling efficiency) is therefore of critical importance for overall biocatalytic activity. Coupling efficiency is also influenced by the concentrations of each of the three components of the CYP multicomponent enzyme (Pennec et al., 2015). White et al. (2017) concluded that CYP153A6 whole cell biohydroxylation was limited by both substrate transmembrane transport and coupling efficiency more than the availability of cofactor.

Several reports have illustrated the advantages of artificial fusion CYP systems such as improved coupling efficiencies, turnover frequencies (TOFs) and solubility compared to natural fusion enzymes or their multicomponent parental enzymes (Honda Malca, 2013, Ciaramella et al., 2017, Bordeaux et al., 2011). Using protein engineering, Kubota et al. (2005) artificially fused the CYP153A13a domain from *Alcanivorax borkumensis SK2* to a flavin mononucleotide/Fe_2S_2-containing reductase (RhFred) from *Rhodococcus sp.* Bordeaux et al. (2011) fully characterised this resultant CYP153A13 fusion construct to illustrate its promising potential for the hydroxylation of medium chain alkanes. The artificial self-sufficient CYP153A13 fusion enzyme is expressed as a single component and consists of electron transfer proteins fused to the heme domain of the CYP, which improves the rate of electron transfer (Refer to Figure 1.1 in Chapter 1.2).

Bordeaux et al. (2011) found that the K_D of CYP153A13 on *n*-octane, *n*-decane and *n*-dodecane were determined to be 29, 16 and 1 μM respectively. These authors concluded that shorter chain alkanes bound less tightly. Furthermore, TOFs obtained for shorter chains were greater than for medium chains, despite the higher binding affinities obtained for medium chain alkanes. Similarly, Nie et al. (2014) found that there is no relationship between binding affinity and CYP153 enzyme activity. The TOF of CYP153A13 on *n*-octane, *n*-decane and *n*-dodecane were 57, 30 and 0.1 min^{-1}, respectively (Bordeaux et al., 2011). The authors provided no speculation on the underlying reasons for this observation. The relationship between lower activity and higher substrate-binding affinity could be because of the substrate binding too tightly to the active site, due to the hydrophobic effect occurring between the highly hydrophobic alkane and active site. This would result in a lower alkane hydroxylation rate because the active site is unavailable to new substrate in the proximal area, either because a substrate molecule remains bound, or the alcohol product remains bound and cannot be released effectively (Cooley et al., 2009, Park, 2022).

2.4.3 Whole cell biocatalysis

Whole cell biocatalysts are preferred to purified enzymes for several reasons including catalyst stability, simplified catalyst preparation and cheaper downstream processing costs (Pfruender et al., 2006). Whole cells re-generate expensive cofactor such as NADPH, and other molecules such as ATP, therefore are preferable to crude, purified or immobilized enzymes (Duetz et al., 2001, Wachtmeister and Rother, 2016). Host cells also regenerate cofactor supply to provide electrons to the active site of the enzyme (Wachtmeister and Rother, 2016). The cell membrane provides protection of the enzyme from shear forces and from reaction conditions in the bioreactor environment that may denature isolated enzymes (Lundemo and Woodley, 2015, Duetz et al., 2001).

The use of whole cell systems, however, introduces additional limiting factors to alkane biohydroxylation such as mass transfer, cellular metabolism and health, substrate uptake across the cellular membrane, reactive oxygen species, stress tolerance to solvents and substrate metabolism (Scheps et al., 2011). In whole cell biocatalysis using *E. coli*, alkane substrates likely diffuse passively across the cellular membrane to reach the expressed alkane hydroxylases in the cytoplasm (Staijen et al., 2000). Alternatively outer membrane proteins facilitate alkane transmembrane transport, which can also be recombinantly expressed in host organisms to improve the rate at which substrate is transported across the membrane (Grant et al., 2014, Hua and Wang, 2014). Contact between the active site of the enzyme and the substrate results in selective oxidation to produce primary alcohols. Staijen et al. (2000) showed that alkane chain lengths of C_6 to C_{12} are transported across the membrane by diffusion. Similarly, Schrewe et al. (2013) stated that alkanes enter whole cells by diffusion with a concentration gradient. White et al. (2017) illustrated that whole cell biocatalytic activity is substantiality constrained by substrate transport across the membrane barrier to the cytosolic enzyme.

Whole cell biocatalytic systems comprise of either growing cells or resting cells (Scheps et al., 2011). Growing cell systems couple cell division and multiplication with biocatalytic production of the product of interest and cells are metabolically active. In resting cell systems, there is no cell division during the reaction and therefore cellular growth and product yield are decoupled. Growing and resting cell systems require an inexpensive carbon source such as glycerol or glucose to provide cells with energy to maintain cellular health and regenerate metabolic equivalents needed for the desired oxido-reductase biocatalytic alkane activation. Growing cells are more resilient against organic solvent toxicity and can maintain higher biocatalytic activities, but the division of carbon resources and energy between bioconversion, cellular growth, and stress-tolerance can be disadvantageous (Olaofe et al., 2013).

Whole cells require specific environmental conditions, including pH and temperature, and operate in an aqueous media (Halling, 2004). Environmental conditions such as oxygen availability and choice of carbon supply impact biocatalytic activity (Gudiminchi et al., 2012, Olaofe et al., 2013). Lack of oxygen may affect cellular health therefore adequate mass transfer must be ensured for both cellular health and biocatalysis. Gudiminchi et al. (2012) illustrated that in small-scale bioreaction mixtures that contain higher CYP concentrations ranging from 0.5 to 0.8 $\mu mol_{P450} \cdot g_{DCW}^{-1}$ oxygen becomes limiting with an increase in electron uncoupling, resulting in lower alcohol product concentrations. Pennec et al. (2015) showed that there was no significant difference in the results obtained for 24 mL amber vials compared to 40 mL amber vials, therefore assumed oxygen was not limiting. In large bioreactors, increased agitation is implemented to improve mass transfer (Meissner, 2013).

Olaofe et al. (2013) illustrated the importance of the relationship between host-physiology, environment and process conditions behind bioprocess efficiency. Process conditions impact host cell performance, physiological states and cellular behaviour in response to environmental changes. Olaofe et al. (2013)

illustrated a 40-fold increase in biocatalyst efficiency of 2.3 $g_{1\text{-octanol}} \cdot g_{dcw}^{-1}$ compared to previous studies by changing environmental factors such as medium composition, carbon source, feeding strategy and product toxicity (Olaofe et al., 2013). Olaofe et al. (2013) concluded that along with metabolic and strain engineering, environmental and process conditions, influences on host physiology, gene expression and biocatalytic activity should also be strongly emphasised in bioprocess design.

In conclusion, optimal recombinant protein expression, function and therefore biocatalytic activity is not only dependent on cellular intrinsic factors, but also on extrinsic factors that promote biomass accumulation and high-density cell growth (Olaofe et al., 2013). Growth rate of *E. coli* is commonly represented by the Monod model (Equation 2.2) (Kovárová-Kovar and Egli, 1998).

$$\mu = \mu_{max} \frac{[S]}{K_s + [S]} \qquad \textit{Equation 2.2}$$

Where μ is the specific rate of biomass growth, μ_{max} is the maximum specific growth rate, K_s substrate saturation constant and $[S]$ is the substrate concentration. This model was developed to describe the impact of rate-limiting substrate such as carbon on microorganism growth rate of an organism. Medium composition plays a critical role in cellular growth rate, which in turn effects protein expression and therefore overall biotransformation efficiency (Olaofe et al., 2013, Gudiminchi et al., 2012).

Luria Bertani (LB) medium is undefined and promotes high biomass accumulation. LB, however, may cause difficulties in regulating growth rate, which results in higher risk for protein misfolding (Makrides, 1996). The carbon source of the medium also plays a role in cofactor regeneration needed for alkane hydroxylase activity (Olaofe et al., 2013, Gudiminchi et al., 2012). The best carbon source choices for *E. coli* are glucose or glycerol. Glucose results in less expensive downstream processing, faster growth rates and higher biocatalyst stability. Excess concentrations of glucose however can result in overflow metabolism and by-product acetic acid formation (Bühler et al., 2003). An intermittent glucose feeding system was shown to prevent overflow metabolism (Olaofe et al., 2013). Glycerol is said to contribute to biocatalyst stability and heterologous protein expression and less acetic acid formation.

Gudiminchi et al. (2012) illustrated that the product formation rate significantly increased with NADH/NAD ratio and NADH supplied via the glycolysis of glucose (Gudiminchi et al., 2012). *E. coli* cells express heme in low concentrations, which is required for the CYP's active site. Therefore the growth medium should also include a heme precursor such as 5-aminolevunilic acid to aid its effective expression (Olaofe et al., 2013). These findings illustrate the importance of medium composition and the influence of these on enzyme functioning.

Stress-related environmental effects on the host cell induces changes in the functioning of the metabolic network, which is responsible for regulating enzyme synthesis and cofactor regeneration (Blank et al., 2008). Product and substrate toxicity causes physiological changes in the membrane, resulting in leakage of ions and cofactors and cell lysis. Product and substrate may also accumulate in the membrane bilayer (White et al., 2017). Non-polar hydrophobic compounds have detrimental effects to cellular health, damaging cellular structures and membranes (Brink et al., 1988, Halling, 1994). Furthermore, exposure of biocatalyst to high concentrations of substrates and products may result in reaction inhibition, biocatalyst toxicity, instability, or deactivation.

Loss of intracellular ions and cofactor disrupts cellular functioning which impacts biocatalytic activity (Grant et al., 2011). Solvent toxicity can be represented by the logarithmic ratio of the partition coefficient ($\log P_{OW}$) in relation to the standard 1:1 octanol to water (Olaofe et al., 2013). Compounds

with $\log P_{OW}$ values in a range of 1 to 4 are toxic, less soluble in water and cause membrane partitioning and disintegration. The relationship between solvent toxicity and biocatalytic activity and metabolism of a recombinant solvent-tolerant strain of *Pseudomonas* was investigated using a systems biology approach (Blank et al., 2008). These authors showed that redox metabolic pathways divide between cellular growth and maintenance, cofactor regeneration and solvent tolerance. NAD(P)H supplied by the glycolysis metabolic pathway from glucose in the media was divided between cellular processes such as cellular growth, biocatalysis and other processes such as cellular maintenance including stress-tolerance.

Blank et al. (2008) showed that the presence of toxic solvents caused a decrease in biomass yield and growth rate while there was an increase in the specific carbon uptake rate. This observation was related to the 'driven by demand' concept first described by Koebmann et al. (2002), which describes how metabolic energy demand of an organism is dependent on the rate at which the metabolic network operates (Koebmann et al., 2002).

Despite the events that cause division of NAD(P)H to different metabolic pathways and cellular functioning, Blank et al. (2008), used a genomic stoichiometric model to show that the presence of bis(2-ethylhexyl)phthalate (BEHP) – a carrier solvent that mitigates solvent toxicity – improves the need of energy for cellular maintenance and thus energy availability to drive the biocatalytic reaction increases (Blank et al., 2008). Blank et al. (2008) and Olaofe et al. (2013) both showed that glucose consumption increased when BEHP was present, compared to the absence of BEHP which illustrated an increase in metabolic capacity for cofactor regeneration for the biohydroxylation reaction. Olaofe et al. (2013) showed that BEHP mitigated product toxicity and product inhibition because cells remained metabolically active for 72 hours longer in the presence of BEHP.

Lower solubility and consequential phase separation of longer chain alcohols decreases toxicity and product extraction costs (Lennen and Pfleger, 2013). It has been previously shown that longer chain alkanes such as *n*-dodecane exhibit less toxicity toward *E. coli* than *n*-octane due to its lower solubility but this also limits substrate access (Grant et al., 2011). This highlights the major trade-off between the concentration that maintains cellular health and substrate availability to the biocatalyst (Leon et al., 1998, Straathof et al., 1992). This is known as the critical concentration. The use of a second carrier phase and product sink allows in situ product removal, which alleviates product toxicity in the aqueous biocatalyst residing phase, known as a 2LPS.

2.5 Reaction conditions and characterisation

2.5.1 2-Liquid-Phase System

Two liquid immiscible phases make up the 2-Liquid-Phase-Systems (2LPSs), consisting of the aqueous reactive phase containing the whole cell biocatalyst and an organic non-reactive phase. Whole cell biocatalysts reside in a buffered aqueous medium for stability. Hydrophobic substrates show limited solubility in aqueous media. Whole cell biotransformation of alkanes into alcohols presents challenges such as product/substrate toxicity and poor water-solubility (Bühler et al., 2003). To this avail, the 2LPS was developed for whole cell biocatalysis of reactions involving highly hydrophobic substrates such as long chain alkanes.

The top organic phase consists of an organic solvent that acts as a substrate reservoir and product sink, thereby tightly regulating the concentration of substrate and products in the aqueous phase (Eckstein et al., 2006b, Kim et al., 2007). Therefore, high concentrations of substrate and product are possible in the overall system without reaction substrate/product inhibition or damage to cellular health

(Bühler et al., 2003, Garikipati et al., 2009, Brink et al., 1988). The 2LPS suffers from limited interfacial substrate and oxygen mass transfer needed for cellular metabolism to whole cell biocatalysts residing in the aqueous medium (Olaofe et al., 2013, Bühler et al., 2003).

The 2LPS incorporates *in situ* product recovery into the alkane bioconversion process in single-unit reaction vessel (Leon et al., 1998). There has recently been an increased preference for heterogenous or multiphase catalysis compared to homogenous catalysis, since it eliminates separation and downstream processing steps to isolate product or recycle catalyst (Kara and Liese, 2019, Keim, 2003). In the case of whole cell biocatalysis, substrate/product transmembrane transport limitations are also introduced. Biocatalytic 2LPSs comprise of three basic components: (1) Enzyme systems located within the cell, (2) Solvent and (3) Substrate/Product. Each of these components interact with one another which contribute to 'apparent' experimental reaction kinetics (Ferrario and Pleiss, 2019). Kinetic parameters drawn from an experimental progress reaction curve are influenced by the following sequential steps:

i. The availability of substrate to aqueous residing biocatalyst. This can be evaluated using partitioning coefficient, effective concentration, thermodynamic activity, or mass transfer.

ii. Transport rate of substrate across the lipid bilayer membrane of the whole cell biocatalyst into the cytoplasm where the CYP153 enzyme resides.

iii. Enzyme-substrate (ES) contact and binding affinity, as well the impact of substrate-solvent interactions on these.

iv. The rate of formation of transition state or activated complex, which describes how fast substrate reacts to form a product in the presence of bound functional enzyme. The biocatalytic hydroxylation mechanism utilises electron carrier proteins to selectively oxidise linear alkane chains at the primary carbon, producing linear primary medium to long-chain alcohols.

v. The rate at which free product is transported through the cellular cytoplasm across the cellular membrane to the immediate surrounding aqueous phase.

vi. The product partitioning, extraction, and mass transfer into the top organic phase from the aqueous phase.

Figure 2.1 illustrates these reaction rate-limiting steps. Steps (i) to (vi) are inter-dependent to obtain a system of overall emergent properties depicted in an experimental progress reaction curve of measured product over time. Some of these steps require thermodynamic, mass transfer and solvation effects analyses, while others are concerned with biology, such as transmembrane substrate transport, ES contact and intrinsic enzyme kinetics. Interactions between these components include solvent-substrate, enzyme-solvent and ES which result in apparent biocatalytic reaction rates (Ferrario and Pleiss, 2019). Crude experimental observations of biocatalytic reaction rates present system-level effects of the thermodynamics, the interfacial and transmembrane mass transfer kinetics as well as the intrinsic enzyme kinetics. Interpretation of biocatalytic progress curve reaction rate data should be done carefully by decoupling ES from solvent-substrate interactions.

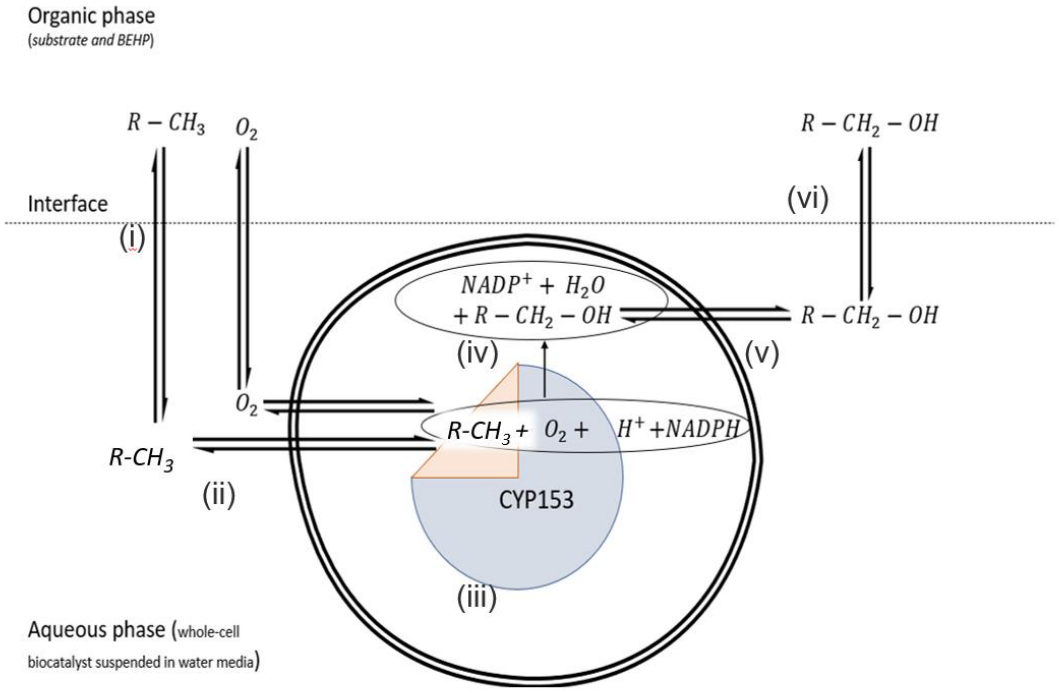


Figure 2.1 An illustration of the potential rate-limiting steps of a whole cell biocatalytic 2LPS.

2.5.2 Biocatalyst-Solvent interaction

Enzymes and whole cell biocatalysts are sensitive to harsh reaction conditions, limiting the range of operation parameters and reaction media chosen. Before the 1970s, water-miscible solvents were used to create monophasic systems in which the solubility of hydrophobic compounds increased to improve ES contact (Halling, 1994). Concentrations of solvents had to be less than 50% (v/v) to avoid biocatalyst inactivation (Pesci, 2017). Low-water and unconventional biocatalytic systems such as 'neat substrate' and 'micro-aqueous' systems are gaining significant attention (Kara and Liese, 2019). Immobilisation can be used with neat substrate systems to improve substrate availability to the biocatalyst (Findeisen et al., 2014). Immobilisation of whole cells facilitate recovery and stability of the catalyst. Whole cells are attached to solid carriers via covalent or adsorptive bonding or entrapment in polymeric matrices (Findeisen et al., 2014).

Pesci et al. (2017) showed that higher substrate concentrations resulted in lower bioconversion due to biocatalyst deactivation and not substrate inhibition (Pesci, 2017). Therefore, it is crucial to consider biocatalyst deactivation from toxic effects when drawing conclusions from biocatalytic reaction data (Pesci, 2017). While it is essential to alleviate the toxic effects of chemical components towards the biocatalyst, it should also be ensured that the concentrations are still maximised to optimise product yield of the biocatalytic reaction.

Water immiscible organic solvents were used from the 1970s onwards, which proved to show substantial improvement in biocatalytic activity (Halling, 1994). This was not expected because solvents of lower polarity were expected to cause higher rates of biocatalyst deactivation. It was soon realised

that the split of the two liquid phases causes less exposure of toxic solvents and substrates to the biocatalyst.

2.5.3 Biocatalyst-substrate interactions

2.5.3.1 Whole cell biocatalyst transmembrane substrate transport

Once the alkane reaches the whole cell biocatalyst residing either at the interface or in the bulk of the aqueous media solution, the cellular membrane forms a barrier to the hydrophobic alkane substrate. This decreases the substrate supply to the cytosolic enzyme system and may decrease reaction rates. Studies have confirmed that substrate transmembrane transport across the cellular membrane is a major limiting factor in whole cell biocatalysis (Julsing et al., 2012, White et al., 2017, Grant et al., 2011).

Several methods have been tested to improve transmembrane substrate transport by either chemical or physical cell membrane permeabilisation (White et al., 2017). *Pseudomonas* species produce their own biosurfactants such as rhamnolipids to increase uptake of hydrophobic substrates into the cellular membrane (Blank et al., 2008, Julsing et al., 2012). Membrane permeability has been increased by adding natural or synthetic surfactants (Triton X-100), cosolvents (dimethyl sulfoxide), chemical permeabilisers and chelating agents (EDTA). Physical methods such as heat shock treatment have also been tested. These methods have shown limited application for long-term and sustainable biohydroxylation because whole cell biocatalysts are vulnerable to cellular damage, causing activity loss and complicated downstream processing (Julsing et al., 2012, White et al., 2017).

Genetic engineering of *E. coli* to express the AlkL transporter outer membrane protein showed a 62-fold increase compared to the microbial host without the AlkL protein, showing a maximal activity of 87 $U \cdot g_{cdw}^{-1}$ (Julsing et al., 2012). The expression of AlkL showed a five-fold improvement of hydroxylation activities of (*S)*-limonene to (*S*)-perrillyl alcohol in a 2LPS (Cornelissen et al., 2013). White et al. (2017) illustrated that chemical permeabilisation improved initial whole cell biohydroxylation rates but compromised biocatalyst stability. It was however found that the mechanical permeabilised cells co-expressing dehydrogenase was superior to whole cells co-expressing dehydrogenase. The mechanical disruption to the cellular membrane improved substrate transmembrane transport but cofactor leaked out of the cell and therefore extra supply was needed from the co-expression of dehydrogenase. An alternative option to increasing the permeability of the outer membrane of *E. coli* is through induced mutations (Ni and Chen, 2004).

Grant et al. (2011) showed that limited substrate transport across the membrane was the most constraining factor on C_5-C_{12} alkane oxidation productivity in a system consisting of engineered *E. coli* whole cells expressing the AlkB complex from *Pseudomonas putida GPo1*. Longer chain alkanes induce recombinant AlkB expression. The bioconversion rate of different chain-length alkanes to primary alcohols and carboxylic acids was measured using a growing cell system. Maximum specific activity on *n*-dodecane was 3.5 $\mu mol_{1-dodecanol} \cdot g_{DCW}^{-1} \cdot min^{-1}$ seven times smaller than on *n*-octane ($21 \mu mol_{1-octanol} \cdot g_{DCW}^{-1} \cdot min^{-1}$), while the volumetric rate was three-fold higher on *n*-dodecane than on *n*-octane. The lower solubility of *n*-dodecane compared to *n*-octane explained this observation. Grant et al. (2011) detected lower expression levels of alkB with *n*-dodecane due to its limited accessibility and transport across the cell membrane and liquid-liquid (LL) interface into the whole cell. Simultaneously, the lower solubility and toxicity of *n*-dodecane compared to *n*-octane reduced the toxicity that cells experience, thereby enabling the recombinant *E. coli* to reach higher cell densities.

Interestingly, Grant et al. (2011) found that 1-dodecanol competes with *n*-dodecane for the enzyme's active site, resulting in the formation of higher levels of dodecanoic acid than the primary alcohol

counterpart. The observed increase in dodecanoic acid production when primary alcohol was added indicated that AlkB preferred 1-dodecanol as a substrate, either due to its lower K_m compared to dodecane or higher accessibility to the enzyme compared to *n*-dodecane because of solubility. The solubility of 1-dodecane ($3.7\times10^{-6} g \cdot L_{water}^{-1}$) is 1000-fold less than that of 1-dodecanol ($4\times10^{-3} g \cdot L_{water}^{-1}$). The access and availability of substrates depend on a combination of substrate aqueous solubility, rate of transmembrane transport, or membrane permeabilisation (van Beilen and Funhoff, 2005, Julsing et al., 2012).

The system showed higher overoxidation with *n*-dodecane than *n*-octane, which calls attention to the competition of substrates of similar nature to occupy and access the enzyme's active site. It was concluded that the most likely rationale was substrate access, indicating that longer chains possess lower solubility and, therefore, limited ability to access the cytoplasmic enzyme active site compared to *n*-octane (Grant et al., 2011). The findings by Grant et al. (2011) highlighted the need to decouple ES from solvent-substrate interactions and to consider substrate solubility when interpreting experimental results and K_m values, further discussed below. Another example provided by Bordeaux et al. (2011) reported that *in vivo* correlates with *in vitro* TOF for linear alkanes. For cyclohexane, however, *in vitro* TOF was lower than *in vivo*. These authors postulated that cyclohexane entered whole cells with greater ease than the linear alkane but could be more toxic (Bordeaux et al., 2011).

2.5.3.2 Intrinsic enzyme kinetics and enzyme-substrate binding affinity

Enzyme kinetics is a branch of study that aims to understand and express factors that influence the rate of enzyme-catalysed reactions. Mathematical equations describe reaction rates and factors that affect these rates (Robinson, 2015). The kinetics of a two-step enzyme catalysed reaction (Equation 2.3) can be characterised by the most widely applied model, the Michaelis-Menten (MM) equation (Equation 2.4). MM kinetics describes reaction rates relative to substrate concentration and initial reaction velocity (v_0). The v_0 of enzyme-catalysed reactions from a series of individual experiments containing different substrate concentrations is measured. In 2LPSs, this approach is complicated to achieve practically because substrate concentration is governed by partitioning across the LL interface. For a general reaction seen in Equation 2.3, the MM model assumes that the binding step (k_1; k_{-1}) is faster than the rate-limiting catalytic step (k_{cat}).

$$E + S \underset{k_{-1}}{\overset{k_1}{\leftrightarrow}} ES \overset{k_{cat}}{\leftrightarrow} E + P \qquad \textit{Equation 2.3}$$

Therefore, the equilibrium ratios of $[E]$, $[S]$, and $[ES]$ are rapidly reached. The catalytic rate (k_{cat}) is a constant representing the maximum number of substrate molecules that can convert into product per molecule per enzyme per unit of time. For simplicity, TOF and k_{cat} are used interchangeably in the field of catalysis (Kozuch and Martin, 2012).

Biocatalytic reactions typically follow mixed-order kinetics (one and zero)(Al-Haque et al., 2012). Lauchnor et al. (2015) found that the first-order relationship to depict whole cell kinetics of ureolysis was more suitable than the MM model, especially within the first stages of the reaction. In the early stages of the reaction, it is assumed that product concentration approximates zero in the early stages of the reaction. The reaction is assumed to reach a steady-state, in which $[ES]$ becomes constant. Substrate concentration is assumed to be significantly larger than the total enzyme concentration, so

the substrate bound in the ES form is negligible (Palmer and Bonner, 2007). An enzyme-catalysed reaction's 'transition state', in which substrate is bound to enzyme to form the ES complex, is characterised by the lowering of free energy difference between the substrate and product i.e., activation energy (Equation 2.3).

The v_0 is defined as the velocity occurring before 10% substrate to product conversion or where the product concentration is increasing linearly over time (Yadav and Magadum, 2017, Palmer and Bonner, 2007). By varying substrate concentration, initial velocity increases hyperbolically until an increase in substrate concentration does not impact the reaction rate, signifying the maximum velocity (v_{max}). The MM constant K_m is the substrate concentration at which half v_{max} is reached (Equation 2.4). At v_{max}, all active sites of enzymes are bound to substrate in a transition-complex state, and the reaction reaches a saturation point.

$$v_0 = \frac{v_{max}[S]}{[S] + K_m} \qquad \textit{Equation 2.4}$$

The rate of product formation or substrate disappearance over time also measures enzyme reaction rates. Product concentration increases as the reaction progresses, while substrate concentration decreases (Figure 2.2). These very first few moments of the reaction are denoted as v_o, calculated by the slope of the tangent of the graph between two carefully selected time points (t_1 and t_0). The v_0 of a reaction is known to be an accurate measure of reaction rate because the amount of substrate added is known, product inhibition does not occur, and no enzyme deactivation has occurred.

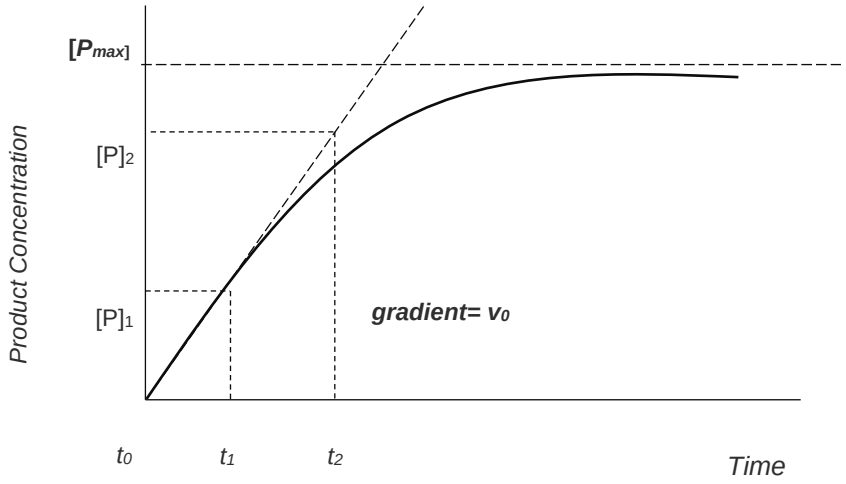


Figure 2.2 The formation of product over time. The initial velocity (v_o) of the reaction is calculated by the gradient of the tangent to the initial curve. The maximum velocity (v_{max}) is the upper limit of the reaction velocity, where the maximum concentration of product is formed (P_{max})

The ES complex allows proximity of the chemical reacting groups and active catalytic sites. Binding constants/dissociation constant of ES (K_D) indicates the affinity an enzyme has for the substrate (Equation 2.5). A smaller K_D value signifies that the enzyme becomes saturated at minimal amounts of substrate and v_{max} is reached at low substrate concentrations, which means the enzyme's affinity for the substrate of interest is strong (Robinson, 2015).

Funhoff et al. (2006) determined that the apparent K_D of CYP153A6 for *n*-octane, *n*-nonane and *n*-decane were 0.17, 0.022, 0.010 μM, respectively. Bordeaux et al. (2011) determined that the K_D of CYP153A13 for *n*-octane and *n*-decane was 29 and 16 μM, respectively. These values were determined

using purified enzymes by recording pure enzyme's UV spectra and comparing them to a sample containing both enzymes and substrate.

$$K_D = \frac{k_{-1}}{k_1} \qquad \textit{Equation 2.5}$$

The difference between K_m and K_D is the presence of k_{cat} in the numerator (Equation 2.6). Under the MM model and rapid equilibrium assumption, the unbinding rate is faster than the catalytic rate of the chemical step (i.e., $k_{-1} \gg k_{cat}$) then $K_m = K_D$, known as the Rapid Equilibrium Approximation (REA). Therefore, substrates with a lower K_m possess a more suitable shape to the enzyme's binding site to improve catalytic efficiency. The K_m^c value assists in determining the appropriate concentration of substrate for a reaction (Tipton et al., 2014). The catalytic efficiency (k_{cat}/K_m) illustrates the frequency of collision between enzyme and substrate molecules and the specificity of an enzyme for a particular substrate. K_m^c values are therefore expectedly larger than the K_D values by the factor k_{cat}/k_1.

$$K_m = \frac{k_{-1} + k_{cat}}{k_1} \qquad \textit{Equation 2.6}$$

The development of molecular simulation models enables insight into ES interactions, substrate docking into enzyme active sites, the solvation effect, the energy involved in ES interaction, and the Gibbs free energy of the reaction's transition state (Braiuca et al., 2006). The simulation of the transition state of a reaction requires the consideration of bond breaking and formation as well as unstable transition state structures, which is too complex for molecular mechanic predictive approaches (Ferrario and Pleiss, 2019). Quantum mechanic (QM) methods can predict and handle transition states well but require expensive computational equipment. Hybrid methods that utilise molecular mechanic and QM predictive tools are attractive since ES interaction can be elucidated using QM, while the rest can be predicted using simpler molecular mechanic approaches (Braiuca et al., 2006).

Accurate ES binding studies require properly folded and functional enzymes. MM model reactions are typically modelled in aqueous media with ideal conditions and aim to determine ‘intrinsic’ kinetic parameters (Palmer and Bonner, 2007, Laane et al., 2009). In classical enzyme kinetic studies with ideal and aqueous conditions, steps (iii) and (iv) from Figure 2.1a are the only steps considered.

The MM model is based on assumptions not applicable in real-life applications such as biocatalytic 2LPSs. K_m is measured experimentally, but the value determined does not reveal the ‘true’ ES binding affinity due to the presence of intervening factors such as insufficient ES contact, biocatalyst deactivation due to compound toxicity, deficits in substrate-solvent and product-solvent interaction, mass transfer limitations, transmembrane transport, and limited substrate partitioning. In an ideal system, v_{max} is reached, and 100% of the substrate is selectively converted to the desired product.

Linear optimisation techniques of the MM model were developed, such as the Lineweaver-Burk plot in 1934 by inverting the MM equation to obtain a linear plot to determine parameters such as v_{max} and K_m by extrapolation (Palmer and Bonner, 2007). There are disadvantages of linear optimisation techniques, such as statistically over emphasising and thereby introducing errors into the parameter estimation at lower substrate concentrations because the scatter of the points around a straight line does not follow Gaussian distribution. Therefore, several studies sought to develop iterative nonlinear optimisation techniques, more easily performed in the age of computational power (Jeričević and Kušter, 2005). Computational power enables the direct fitting of product concentration as a function of time to the MM model equation.

Experimental data can be differentiated to obtain rates, or the kinetic model is integrated to obtain product concentration as a function of time data (Schwert, 1969, Goličnik, 2013, Duggleby, 1995). There are several advantages to studying complete reaction progress curves of product concentration as a function of time data (Goličnik, 2013, Duggleby, 1995). Progress reaction curves provide insight into the dependence of enzymes on substrate concentration and product concentration. Moreover, progress reaction curve kinetic analyses illustrate information about the complete rate equation (Choi et al., 2017, Schwert, 1969). In the past, the complexity of the data analysis of complete reaction progress curves has hindered biologists from analysing the kinetic behaviour of enzymes. Mathematical tools have been made available in the past decade to make data analysis and handling easier.

Nonlinear regression (NLR) techniques can be employed to estimate kinetic parameters (Al-Haque et al., 2012). In biocatalysis, NLR analyses minimises the differences between the MM model prediction and corresponding experimental data collected as product concentration at several time points. The optimisation technique commonly used is the minimisation of the sum of the differences between the model and experimental values. NLR techniques require computational software to fit the experimental data to the model as well as good initial estimates (Jeričević and Kušter, 2005). Poor initial estimates reduce the chance of convergence to the global minimum.

While the 2LPS offers several advantages, this reaction system is complex, and its behaviour is more difficult to understand than usual biocatalysis performed in aqueous media. In 2LPSs, the substrate concentration depends on substrate partitioning behaviour. For example, suppose the substrate partitioning into the aqueous reactive phase occurs to a larger extent than substrate-to-product bioconversion. In that case, one could conclude that the reaction is limited by intrinsic kinetics and not by mass transfer or thermodynamic-regulated partitioning.

2.5.4 Substrate/Product-Solvent interaction and availability of substrate to biocatalyst

As the reaction proceeds, substrate and product surround the whole cell biocatalyst in the aqueous phase. Substrate and product concentrations in the aqueous phase are governed by interfacial mass transfer, differential solvation and partitioning of substrates/products (Straathof, 2003, Halling, 1994, Eckstein et al., 2006b). Thermodynamic and kinetic limitations in 2LPSs impact the extent of the reaction and its rate. Catalytic reactions can be kinetically or thermodynamically controlled, which determines which products are produced. In alkane hydroxylation whole cell biocatalytic 2LPSs, the kinetics specifically refers to the rate of the reaction catalysed by CYP153 for the biohydroxylation of alkanes, the reaction pathway and even the mass transfer. On the other hand, low alkane aqueous solubility, substrate partitioning and phase equilibrium in 2LPSs constitute the thermodynamics of the reaction process. To illustrate the difference between thermodynamics and kinetics more clearly, the two different reactions in Equation 2.7 and 2.8 have the same equilibrium constants (K_{eq}).

$$CCl_4 + H_2O \leftrightarrow CO_2 + 2HCl \qquad K_{eq} = 5 \cdot 10^{40}$$

Equation 2.7

$$SiCl_2 + H_2O \leftrightarrow SiO_2 + HCl \qquad K_{eq} = 5 \cdot 10^{40}$$

Equation 2.8

While both chemical reactions have similar equilibrium constants, the reaction in Equation 2.8 occurs immediately to yield product in contrast to reaction in Equation 2.7 that occurs only after a substantial

amount of time (Schmitz, 2012). These two reactions therefore have similar thermodynamics in terms of equilibrium constants but occur at vastly different rates and therefore their kinetics are not similar. The relevancy and application of these to alkane hydroxylation utilising whole cell biocatalytic 2LPSs will be discussed further in Chapter 2.5.4.1 and Chapter 2.5.4.2 below.

Both thermodynamics and kinetics can be manipulated to improve the rate (how rapidly product forms) and yield of a reaction (the extent to which product forms) (Sandoval et al., 2001, Rezai and Traa, 2008, van Tol et al., 1995, Nakanishi et al., 1986, Nekrasov et al., 2018). The definitions of yield and rate of a biocatalytic reaction are provided in Table 2.2 but clarified further according to 'The Paraffin Activation' sector apart of c*change Centre of Excellence in Catalysis Research found at https://www.cchange.ac.za/paraffin-activation/. The yield is defined by the amount of desired product formed per unit amount of substrate consumed or by specific product concentration known as turnover number (TON), which is the amount of product formed per unit amount of enzyme. Space-time yield or volumetric rate is the amount of product formed per unit of working volume per unit of time. The rate is alternatively defined by the turnover frequency (TOF), which measures the amount of product formed per unit amount of enzyme per unit of time.

The productivity of a biocatalytic 2LPS and substrate availability to the whole cell biocatalyst in the aqueous phase depends on: (1) the partition coefficient, which is a thermodynamic-related variable, and (2) a mass transfer coefficient, which is a kinetic-related variable (Kara and Liese, 2019). Thermodynamically regulated partitioning governs the concentration of substrate across the two-liquid phases and, thereby, the maximum amount of substrate available to the biocatalyst in the aqueous phase (Keim, 2003).

On the other hand, the enzyme-catalysed reaction is kinetically controlled by the mechanistic events, reaction pathways, and in 2LPSs, mass transfer across the LL interface (Straathof, 2003). Mass transfer governs the rate at which substrate is transferred across the LL boundary during the reaction. Mechanistic events include the binding, dissociation, and catalytic rate-limiting steps, represented by constants k_1, k_2 and k_{cat} (Kara and Liese, 2019).

Straathof et al. (2003) evaluated the factors that affected the reaction kinetics of biocatalytic 2LPSs and stated that reaction rates are affected by substrate concentration, initial enzyme concentration, and volume-specific interfacial area. Straathof et al. (2003) illustrated that these effects are linear at lower product concentrations, while at higher product concentrations, they will reach saturation. The reaction's location is also impacted by thermodynamics and mass transfer kinetics. The biotransformation reaction occurs exclusively in the aqueous phase, either in the bulk solution or at the organic-aqueous liquid interface (Straathof, 2003). Since longer-chain alkanes exhibit poor water solubility, the alkane concentration is higher at the interface than in the bulk aqueous solution.

The odd-even effect of liquid alkanes could impact the kinetics and thermodynamics of a 2LPS due to its impact on physicochemical properties (Dhiman et al., 2022). The odd-even effect of liquid alkanes refers to the alternation of physiochemical properties in relation to the number of carbon atoms. For example, the odd-even effect causes even-chain carbon lengths to have higher melting temperatures and densities than odd-chain carbon lengths (Yang et al., 2016, Braun et al., 2013).

2.5.4.1 Mass transfer

The mass transfer limitations in biocatalytic alkane oxidation 2LPSs arise when the alkane's mass transfer rate from the organic phase to the aqueous phase is slower than the oxidation reaction rate (Straathof, 2003). The diffusion of substrate molecules into the aqueous phase is impacted by solvent-

substrate molecular interactions, substrate molecule size, and solvent viscosity (Ferrario and Pleiss, 2019, Celebi et al., 2020).

Effective mass transfer across the LL interface relies on a large interfacial area and sufficient agitation. Studies have shown a positively correlated relationship between reaction rate and interfacial surface area. Interestingly, biocatalyst cellular concentration contributing to the viscosity of the aqueous phase and cell clustering at the LL interface both play a role in mass transfer (Kara and Liese, 2019, Straathof, 2003, Brink et al., 1988). In biohydroxylation systems using *E. coli*, the alkane substrate is transferred into the aqueous medium via solvent droplets. The same mechanism occurs for the mass transport of the product from the aqueous to the organic phase (Olaofe et al., 2013).

The solubility of the substrate is a significant factor in the rate at which this transfer occurs. For example, *n*-octane has a solubility of 7 mg ·L^{-1} at 20°C, and the transfer rates range from 10 to 60 $mmol \cdot L^{-1}h^{-1}$ in a large-scale bioreactor, depending on the power input controlling the agitation (Olaofe et al., 2013). The addition of chemical constituents such as dimethyl-sulfoxide (DMSO) and Triton X-100 promotes substrate solubility, phase-mixing, and mass transfer. Alkanes are highly hydrophobic molecules that tend to collapse in the presence of surrounding polar molecules such as water (Ferguson et al., 2009). Process design and additional solvents or biosurfactants reduce this hydrophobic collapse, increase substrate solubility in the aqueous phase, and increase overall biocatalytic activity (Grant et al., 2011, Ferguson et al., 2009).

Diffusion coefficients can be calculated and predicted using hydrodynamic models such as the Stokes-Einstein relation or molecular dynamic (MD) or Kinetic Monte Carlo simulations. Diffusion coefficients and MD methods can also be used to study the net mass transfer of molecules from one phase to another (Celebi et al., 2020). Straathof et al. (2003) developed a mass transfer model to describe the mass transfer of a substrate from the organic phase to the aqueous phase. According to Straathof et al. (2003), the depletion of a substrate over time in the organic phase (c_A^{org}) equates to the reaction rate of the biocatalysed reaction in the aqueous phase (r_A) multiplied by the volumetric ratio between the aqueous and organic phases (V^{aq}/V^{org}) (Equation 2.9). In this model, it was assumed that the aqueous phase substrate concentration equated to its depletion in the organic phase (Equation 2.9).

$$-\frac{dc_A^{org}}{dt} = \frac{V^{aq}}{V^{org}} r_A \qquad [mol \cdot m^{-3} \cdot s^{-1}] \qquad \textit{Equation 2.9}$$

The reaction rate (r_A) was incorporated into the MM model, given that the reaction rate was dependent on the biocatalyst enzyme concentration in the aqueous phase (c_{EO}^{aq}) as well as the aqueous phase alkane concentration (c_A^{aq}) (Equation 2.10).

$$r_A = \frac{k_{cat} \cdot c_A^{aq}}{K_m + c_A^{aq}} \cdot c_{EO}^{aq} \qquad [mol \cdot m^{-3} \cdot s^{-1}] \qquad \textit{Equation 2.10}$$

The substrate aqueous phase concentration depletion was assumed to be negligible relative to the reaction and mass transfer rates. Therefore, reaction and mass transfer rates could be equated to present a steady state for the aqueous phase (Equation 2.11).

$$k_L a \cdot \left(\frac{c_A^{org}}{m} - c_A^{aq}\right) = \frac{k_{cat} \cdot c_A^{aq}}{K_m + c_A^{aq}} \cdot c_{EO}^{aq} \qquad [mol \cdot m^{-3} \cdot s^{-1}] \qquad \textit{Equation 2.11}$$

where k_L denotes the mass transfer coefficient, a denotes the volume-specific interfacial area in m^2 per m^3 aqueous phase, and m denote the partition coefficient. Therefore, if K_m is smaller than c_A^{aq}, then Equation 2.12 can be used to calculate aqueous phase concentration.

$$c_A^{aq} = \frac{-b_1 + \sqrt{(b_1)^2 + 4\,K_m \cdot c_A^{org}/m}}{2} \qquad [mol \cdot m^{-3}] \qquad \textit{Equation 2.12}$$

where

$$b_1 = \frac{k_{cat} \cdot c_{EO}^{aq}}{k_L a} + K_m - \frac{c_A^{org}}{m} \qquad [mol \cdot m^{-3}] \qquad \textit{Equation 2.13}$$

Using these model equations, Straathof et al. (2003) illustrated that as enzyme and substrate concentration increase, there is a point at which mass transfer becomes limiting. The importance of the partitioning behaviour of substrates in the systems reaction kinetics was highlighted. Given adequate substrate partitioning, these authors investigated whether the reaction occurs at the aqueous-organic liquid interface or in the bulk aqueous solution.

Interfacial reactivity may be higher than reactivity in the bulk aqueous phase because of increased biocatalyst-substrate contact, either due to higher substrate concentrations at the interface due to decreased mass transfer limitation, biocatalyst interfacial absorption or biocatalyst interfacial activation (Brink et al., 1988, Straathof, 2003). Partitioning behaviour affects the reaction kinetics but is rooted in the principles of thermodynamics.

2.5.4.2 Thermodynamics

The thermodynamic behaviour of biocatalytic 2LPSs has been shown to impact process performance significantly (Pesci, 2017). Thermodynamic phenomena indirectly affect biocatalytic activity and specificity. Equilibrium is influenced by various parameters, such as changes in concentration of substrate and products, differential solvation, Gibbs free energies, equilibrium constants, partitioning behaviour, thermodynamic activity (a), solubility and chemical potential of reactant and product components (Eckstein et al., 2006a, Halling, 1994). The evaluation of the thermodynamic behaviour of biocatalytic 2LPSs could provide insight into reaction kinetics and rates and estimation of achievable product yield (Voges et al., 2017, Leon et al., 1998, Halling, 1994).

Several thermodynamic studies focus on 2LPS organic solvent screening to maximise the substrate concentration available to the biocatalyst (Halling, 1994, Halling, 1990). The extent to which nonpolar substrate dissolves in the aqueous phase governs its maximum concentration in the immediate environment of the biocatalyst. The assessment of reaction conditions, substrate supply, and thermodynamic equilibria are matters of process design. Process design and properties such as reaction conditions and choice of carrier solvent all contribute to the equilibrium position of the biocatalytic reaction which influences overall product yield and even the reaction kinetics.

Chemical equilibrium position and energetics of substrates and products impact the product yield biocatalytic reactions (Voges et al., 2017, Janssen et al., 1995). The Gibbs free energy (ΔG) function of a reaction is defined as the energy that is available to do work as the reaction proceeds toward chemical equilibrium. Biocatalysts do not change the position of chemical equilibrium but instead lower the activation energy required to proceed forward, allowing the reaction to reach equilibrium faster. Therefore, the ΔG of the transition state (i.e., the ES complex) of a reaction contributes to the reaction kinetics. Chemical equilibrium constants (K_{eq}) are governed and well understood in terms of Gibbs free energy. The standard state free energy change (ΔG°) relates to the equilibrium constant by Equation 2.14.

$$\Delta G^{\circ} = -RT\ ln\, K_{eq} \qquad \textit{Equation 2.14}$$

where K_{eq} is the equilibrium constant, R is the universal gas constant, and T is the temperature. More specifically, free energy is a state function that expresses the spontaneity of a chemical process in terms of the enthalpy and entropy change of the system. $\Delta G^{\circ\prime}$ and K_{eq} are dependent on molecular structure and stoichiometric coefficients of chemical species involved in the reaction (Mavrovouniotis, 1991). The overall Gibbs energy of formation depends on the Gibbs energy of formation of both reactants and products of a reaction (Equation 2.15).

$$\Delta G^{\circ\prime} = \sum_{i} v_i \Delta\, G_i^{\circ\prime} \qquad \textit{Equation 2.15}$$

where $\Delta\, G_i^{\circ\prime}$is the standard Gibbs free energy of formation of a compound and v_i is the stoichiometric coefficient of a chemical species involved in the reaction. The $\Delta\, G_i^{\circ\prime}$of compounds, therefore, can be used to determine the Gibbs free energy change of a biotransformation reaction (ΔG°) as well as the equilibrium constant K_{eq}. Equation 2.16 is considered for general reaction.

$$A + B \leftrightarrow C + D \qquad \textit{Equation 2.16}$$

The K_{eq} of this reaction is dependent on the relative ratio of the product concentration to the substrate concentrations involved in the reaction (Equation 2.17).

$$K_{eq} = \frac{[C][D]}{[A][B]} \qquad \textit{Equation 2.17}$$

The following equilibria are considered for the two-phase liquid biocatalytic alkane hydroxylation system (Ivanov et al., 1996). $Krxn$ is the chemical equilibrium constant for the reaction (Equation 2.18).

$$R - CH_3 + O_2 + H^+ + NADPH \overset{Krxn}{\longleftrightarrow} R - CH_2 - OH + NADP^+ + H_2O \qquad \textit{Equation 2.18}$$

Several studies have aimed toward maximising 2LPS biocatalytic activity by shifting phase and chemical equilibrium toward the primary alcohol product (Eckstein et al., 2006a, Nakanishi and Matsuno, 1986). Eggers et al. (1989) shifted the chemical equilibrium of a reaction with water as a by-product by adding a water-miscible and organic solvent-immiscible component, glycerol, to the aqueous phase (Eggers et al., 1989). Adding glycerol decreased the water activity or concentration perceived by the whole cell biocatalyst, shifting the reaction equilibria towards product and water by-product production. It is noteworthy that this approach, however, may result in changes in a of other chemical components present in the system.

Ivanov et al. (1996) calculated the Gibbs free energy for both chemical and phase equilibria, respectively, and found that the total Gibbs free energy of the system was governed mainly by the phase equilibria. The net thermodynamic equilibrium of chemical and phase equilibria of a 2LPS biocatalytic system was mainly governed by the partitioning behaviour of substrates and products, while the effects of the reaction equilibria were negligible (Ivanov et al., 1996).

Phase equilibrium describes the partitioning behaviour of chemical substrate and product components between the two immiscible liquid phases (Prpich and Daugulis, 2004, Fam and Daugulis, 2012, Kara and Liese, 2019, Ivanov et al., 1996). The behaviour of chemical components in a 2LPS can be analysed in terms of relative concentrations in the bulk organic and aqueous phases, respectively. When a compound exists in only one form, distribution ratios, and partition coefficients (P_x) are

equivalent (Harvey, 2020). K_p, the partition coefficient, illustrates the affinity of substrate towards the two immiscible liquid phases, respectively (Equation 2.19).

$$R - CH_3(org) \overset{K_p}{\leftrightarrow} R - CH_3(aq) \qquad \textit{Equation 2.19}$$

According to the law of distribution, component *i* is distributed between two immiscible liquids at equilibrium. Thermodynamic equilibrium positions are investigated in terms of concentrations of each component in each phase. For the reaction in Equation 2.19, partition coefficient K_p is used to describe the ratio of the concentration (Equation 2.20) or ratio of mole fraction (Equation 2.21) of the solute (i) of interest in the nonreactive organic phase (α) relative to the aqueous reactive phase (β).

$$K_p = \frac{[i]^{\alpha}}{[i]^{\beta}} \qquad \textit{Equation 2.20}$$

$$K_{p(x)} = \frac{x_{\alpha}}{x_{\beta}} \qquad \textit{Equation 2.21}$$

Phase equilibria can also be shifted to improve the reaction yield. For example, a second organic immiscible liquid phase shifts the phase equilibria by acting as a substrate reservoir or an *in-situ* product extractant (Kim et al., 2007). The Gibbs function can be used to determine the "escaping tendency" of a reaction component in terms of a concentration of a solute (Straathof et al., 1992). The ΔG value of a given reaction can be used to determine the equilibrium constant and the partitioning behaviour of a chemical component in a 2LPS.

Optimisation of a biocatalytic reaction could entail increasing substrate mass transfer and attaining favourable equilibrium concentrations of both substrate and products among the two liquid phases (Fam and Daugulis, 2012). Traditional organic chemistry approaches to shifting equilibria involve changing factors such as temperature, pH, and pressure. Biocatalysts, however, function optimally in a limited range of conditions (Tufvesson et al., 2012).

In batch reaction systems, selecting a water-immiscible organic solvent allows fine-tuning an equilibrium position, in situ product extraction, and substrate supply (Wiltschi et al., 2020). Some studies compared the thermodynamic equilibrium position of systems with different organic solvents (Janssen et al., 1995). Therefore, shifting the equilibrium to favour a higher product yield requires alternative approaches such as organic carrier phase screening and selection, altering phase ratio, or initial substrate ratio (Eckstein et al., 2006a). For example, when the product has a higher affinity for the organic solvent, *in situ* product extraction occurs, removing the product from the immediate environment of the biocatalyst, causing a shift in the thermodynamic equilibrium of the reaction.

There is considerable literature on organic solvent screening because strong evidence indicates its influence on reaction kinetics (van Tol et al., 1995, Laane et al., 2009). After all, the nature of the organic solvent plays a role in the thermodynamic equilibrium position, chemical component partitioning behaviour and solubility, biocatalyst stability, selectivity and activity, product yield and recovery (Halling, 2004, Janssen et al., 1995, Heinemann et al., 2003, Brink et al., 1988). One of the biggest challenges in non-aqueous medium and 2LPSs is the elucidation of the relationship between reaction medium, physicochemical properties, and enzyme activity (Laane et al., 2009). Tufvesson et al. (2014) coupled biocatalyst kinetic and substrate physicochemical property models to optimise and shift equilibrium. Well-justified findings indicated that the substrate's equilibrium constant and physicochemical behaviour were essential for process performance optimisation in biocatalytic 2LPS.

Some other criteria have been established for guidance for the selection of an organic solvent by physicochemical properties such as viscosity, dielectric constants, $\log P_{OW}$ values, polarity, reactivity, toxicity, hydrophobicity, component solvation and partitioning behaviour, and biocompatibility (van Tol et al., 1995). The biocompatible product sinks generally have a $\log P_{OW}$ higher than 4-5 (Schewe et al., 2009)**.** Previous studies show a relationship between solvent properties and biocatalyst activity and stability (Barberis et al., 2006). Solvent selection and solvent-substrate interaction play a significant role in biocatalytic activity. A biocatalytic 2LPS has two solvents: the organic carrier nonreactive and aqueous-based. The equilibrium position of the system, dependent on solvation effects, regulates the concentration of substrates available in the microenvironment of the biocatalyst and, consequently, impacts the substrate supply kinetics and, thereby, reaction kinetics and product yield (Tufvesson et al., 2012, Kim et al., 2007, Heinemann et al., 2003).

At thermodynamic phase equilibrium, the net distribution rate of substrate and product across the LL interface reaches zero (Heinemann et al., 2003). Thermodynamic equilibrium is constantly disrupted because the substrate is consumed and the product released, re-establishing a new equilibrium (Halling, 1994). Despite the constant disruption due to the change in concentration of the chemical species, the 2LPS will always tend towards equilibrium via mass partitioning according to the universal laws of thermodynamics (Halling, 1994, Abu and Woodley, 2015, Grosch et al., 2017). In a biocatalytic 2LPS, as the substrate is bio-converted into the product in the aqueous phase, thermodynamic equilibrium is maintained by replenishing the substrate in the aqueous phase.

Cellular metabolic demand and the thermodynamic equilibrium are the 'driving forces' of substrate partitioning. These driving forces impact substrate availability to the biocatalyst, which creates a 'demand-based' delivery system (Kim et al., 2007, Fam and Daugulis, 2012). Partitioning the alkane into the aqueous phase should permit a concentration that allows availability but is not toxic to the whole cell biocatalyst. Simultaneously, partitioning the resultant product into the organic phase (i.e., product sink) should not allow product inhibition and toxicity to the whole cell biocatalyst.

Partitioning behaviour is governed by a thermodynamic function known as chemical potential (μ). As the reaction proceeds, the chemical potential gradients of substrate and product are created in the 2LPS. These gradients directly affect the rate at which the biocatalytic hydroxylation reaction can occur. Differences in μ between two liquid immiscible phases drive the transfer of component i from one phase into another. At equilibrium, the μ is identical in both liquid phases, and free Gibbs energy is minimised. An alkane-water mixture reaches thermodynamic equilibrium at constant pressure and temperature when the lowest Gibbs free energy state is reached within the system. The chemical potential was defined by Josiah Willard Gibbs more than 140 years ago as the chemical energy (U_c) of one mole of substance (n) (Chen, 2019) (Equation 2.22)

$$\mu = G_m = \frac{U_c}{n} \qquad \textit{Equation 2.22}$$

where G_m denotes the molar Gibbs energy. Component i molecules will transfer into a phase with low chemical potential. In terms of chemical potential, the minimal total Gibbs energy of a system at equilibrium is shown by Equation 2.23.

$$G_{total} = \sum_i G_i n_i \ \ where\ G_i = \mu_i \qquad \textit{Equation 2.23}$$

At equilibrium, μ of a component is identical in phase α and phase β (Equation 2.24).

$$\mu_i^{\alpha} = \mu_i^{\beta} = \cdots \mu_i^{\pi} \quad i = 1,2 \ldots N \qquad \textit{Equation 2.24}$$

where N is the number of chemical components present, and π represents the distinct phases. As stated previously, as the biocatalytic reaction proceeds, substrate consumption and product formation change their concentrations in the aqueous phase of the biocatalysts, and therefore the equilibrium changes (Prpich and Daugulis, 2004). In other words, equilibrium change occurs due to differential solvation.

Chemical potential relates Gibbs free energy to the transfer or motion behaviour of molecules and can be defined by thermodynamic activity (a) (Ferrario and Pleiss, 2019). Thermodynamic activity measures a molecular species' chemical potential or effective concentration in solution for nonideal conditions, while concentration shows the chemical potential in ideal conditions (Halling, 2004) (Equation 2.25). For aqueous-organic 2LPSs at thermodynamic equilibrium, chemical component i, distributed between organic phase A and aqueous phase B thermodynamic activities are assumed to be equal in both phases at equilibrium (Halling, 1994, Walter, 2012).

$$a_A = a_B \qquad \textit{Equation 2.25}$$

Classical enzyme kinetic studies analyse reaction rates occurring in aqueous media under ideal and dilute conditions. In these conditions, it would be correct to assume that concentrations and thermodynamic activities are well correlated and that activity coefficients are close to unity (Voges et al., 2017, Ivanov et al., 1996). On the other hand, thermodynamic analyses allow for the investigation of enzymatic reaction rates in more complex systems such as 2LPSs. In systems with dissimilar molecules, such as aqueous media and long-chain alkanes, molecular properties, such as energy and size asymmetry, differ between the two respective liquid phases, introducing non-ideality. The aqueous phase consists of polar molecules and exhibits different intermolecular interactions, sizes, and shapes than the non-polar hydrocarbon phase.

The deviation of a compound i from ideality can be described by the activity coefficient (γ_i), by reference to the pure liquid compound at the same temperature and pressure (Halling, 1994, Abrams and Prausnitz, 1975, Ferrario et al., 2018). The greater the γ_i, the more nonideal the mixture (Equation 2.26). The product of the component's mole fraction (x_i) and activity coefficient (γ_i) equates to the component's thermodynamic activity (Equation 2.26).

$$a_i = \gamma_i \cdot x_i \qquad \textit{Equation 2.26}$$

Mutual solubilities of components in each of the phases in an LL system depend on temperature and pressure. Activity coefficients larger than unity imply high repulsion between the unlike molecules in the system, which increases average distance and phase separation. Not only does the addition of cosolvents such bis(2-ethylhexyl) phthalate (BEHP) alleviate the substrate and product toxicity to the cell, but it also increases the total number of interactions between similar chemical functional groups of all the different molecules present in the multicomponent mixture. The phenomenon is known as the 'cosolvent effect,' which decreases the free energy needed for the alkanes to partition into the aqueous phase, enhancing the alkanes solubility (Lewandowski, 2008).

The conditions of industrial 2LPS biocatalytic applications are nonideal. Therefore, it is more appropriate to express equilibrium constants that are activity coefficient-based rather than concentration based to

account for nonideality and inter-component molecular interaction. According to Jansen et al. (1995), activities instead of concentrations can be used to predict equilibrium positions under nonideal and non-dilute conditions (Halling, 1994, Janssen et al., 1995). In 2LPSs, the thermodynamic activity of water is close to unity since the aqueous reactive phase is completely immiscible with the organic solvent. According to Ferrario et al. (2018), the activity coefficient can be determined at infinite dilution for compounds of poor aqueous solubility, such as alkanes. For phase equilibrium of a chemical component i, Equations 2.27 and 2.28 can be respectively applied to calculate a_i^r and activity-based solvent-independent equilibrium constant K_{eq}^a.

$$(a_i^r)_{aq} = (a_i^r)_{org} \leftrightarrow (x_i)_{aq} \cdot (\gamma_i^r)_{aq} = (x_i)_{org} \cdot (\gamma_i^r)_{org} \qquad \textit{Equation 2.27}$$

As previously discussed, non-dilute system models are preferred for large-scale, technical, and economic feasibility since this requires high product concentrations. Non-dilute model systems are more suitable for biocatalytic 2LPSs since high product concentrations are desired, and substrates are often loaded at high concentrations. In contrast, Ferrario et al. (2018) stated that activity coefficients of poor aqueous solubility components are independent of concentration and can be computationally calculated at infinite dilution. Either way, the standard state of calculations should be carefully selected and consistent. Activity coefficients can be normalised to Raoult's law, indicated by superscript r, chosen as the standard state for non-dilute solutions. For Raoult's law, $x_i \to 1, \gamma_i \to 1$ (Halling, 1994) (Equation 2.28).

$$K_{eq}^a = \frac{a_C \cdot a_D}{a_A \cdot a_B} = \frac{x_C \cdot x_D}{x_A \cdot x_B} \cdot \frac{\gamma_C^r \cdot \gamma_D^r}{\gamma_A^r \cdot \gamma_B^r} \qquad \textit{Equation 2.28}$$

Thermodynamic 2LPS behaviour analyses allow the characterisation of the limitations that partitioning and substrate and product solubility place on the overall enzyme reaction kinetics (Janssen et al., 1995, Eckstein et al., 2006a). Predictable thermodynamic effects of reaction conditions and physicochemical properties can inform experimental data observations on apparent biocatalytic activity or kinetic parameters. Several models have been developed to predict the thermodynamically dependent behaviours and equilibrium position of aqueous-organic 2LPSs.

Experimental approaches measure substrate and product component concentrations in each organic and aqueous phase under varying reaction conditions such as temperature, pH, and initial substrate concentrations. Studies state that laboratory-based organic solvent screening to optimise the shift in phase equilibria is resource-burdening and tedious. Moreover, complex reaction conditions of biocatalytic 2LPSs introduce difficulty in the experimental determination of activity coefficients. As a result, mathematical equation derivations and predictive models have been developed (Kara and Liese, 2019).

Schmid et al. (1998) derived a theoretical calculation to estimate alkane aqueous phase concentration (c_{oct}^{aq}) in 2LPSs (Equation 2.29). This estimation was based on the *n*-octane molar fraction (x_{oct}^{org}) in the organic phase. The organic phase consisted of *n*-octane and cosolvent 1-hexadecene, which formed an ideal solution; therefore, $\gamma_{octane} = 0.99$ in 1-hexadecene at infinite dilution. The *n*-octane partitioning coefficient for a pure-water-octane system is $P_x = 1 \times 10^7$ and c_{H2O} is the mol-based concentration of pure liquid water 55.5 M (Kertes, 1989). By assuming that the partitioning coefficient is constant over a range of *n*-octane concentrations, Schmid et al. (1998) validated Equation 2.29 to estimate equilibrium *n*-octane concentrations in the aqueous phase of a 2LPS. The aqueous phase concentrations of *n*-

*o*ctane at equilibrium were determined to be between 0.08 μM and 0.22 μM, keeping in mind that *n*-octane has a water solubility of 5 μM at 25˚C (Schmid et al., 1998).

$$c_{oct}^{aq} = \frac{x_{oct}^{org}}{P_x} \cdot c_{H2O} = 5{,}5 \times 10^{-6} \cdot x_{oct}^{org} \qquad \textit{Equation 2.29}$$

Alternatively, activity coefficients can be calculated using thermodynamic calculations, molecular dynamic free energy simulations, equation-of-state molecules, or predicted using group contribution methods (Fredenslund et al., 1975, Abraham et al., 2017, Lewandowski, 2008). These mathematically derived models and expressions provide more information on solvent-substrate interactions and solvation effects, such as substrate availability to the biocatalyst in the 2LPS. Group Contribution Methods (GCMs) are commonly used to predict the thermodynamic behaviour of biocatalytic 2LPSs. Partitioning behaviour and free energy transfer of a molecule between two distinct phases can be predicted by GCMs.

GCMs are based on the local composition models and allow for estimating the thermodynamic properties of compounds. In 1964, the first local composition model (the Wilson equation) was derived (Equation 2.30). Local composition models account for the deviation of local compositions from bulk solution by considering non-random molecular interactions based on molecular size and intermolecular forces located in unique local compositions within the overall reaction mixture (Hristova et al., 2015, Abrams and Prausnitz, 1975). The Wilson equation for 2LPS multicomponent systems is described by excess Gibbs energy (Equation 2.30).

$$\frac{G^E}{RT} = -\sum_i x_i \ln\left(\sum_j x_j \Lambda_{ij}\right) \qquad \textit{Equation 2.30}$$

where x_i denotes the mole fraction of chemical species i, and j is a second chemical species. The Wilson parameter Λ_{ij} is defined by molar volumes of components 1 and 2, denoted by V_1 and V_2 respectively and A represents the interaction energy parameter between two components (Hristova et al., 2015) (Equation 2.31).

$$\Lambda_{ij} = \frac{V_2}{V_1} exp\left(\frac{-A_{12}}{RT}\right) \qquad \textit{Equation 2.31}$$

Experimentally determined activity coefficients of chemical components are used to characterise binary interaction parameters between functional group pairs, which are applied to predict activity coefficients of components in systems that lack phase equilibrium experimental data (Fredenslund et al., 1975). GCMs are advantageous when experimental phase equilibrium data are unavailable. The calculation of activity coefficients by GCMs employs the surface area, size, and volume of each functional group, the number of occurrences of that functional group in the molecule, and the interaction energy between the functional groups in the solution. GCMs, subdivide chemical compounds into functional and chemical groups (Klamt, 2005). The additive effects of these functional groups contribute to the overall thermodynamic property of interest (Mavrovouniotis, 1990).

Previously mentioned $\log P_{OW}$ is the most developed and commonly employed GCM in applications such as drug design (Marrero and Gani, 2002). Models based on $\log P_{OW}$ require that the solute of interest be thermodynamically studied, while keeping temperature, pressure, and solvent phase constant (Wienke and Gmehling, 1998). This approach is unsuitable for systems with dynamic variable compositional changes, such as the biocatalytic 2LPS.

The UNIFAC model, commonly used in 2LPS biocatalysis, derived from the UNIversal QUAsi-Chemical (UNIQUAC), can predict to liquid-phase activity coefficients of chemical components in a nonelectrolyte solution (Fredenslund et al., 1975). Several biocatalytic studies apply the UNIFAC model to predict the partitioning behaviour of chemical species between an aqueous bottom phase and an organic carrier solvent top phase (Smith, 1950, Lewandowski, 2008, Jakob et al., 2006, Navas et al., 2010, Tufvesson et al., 2012, Mavrovouniotis, 1990).

The UNIFAC model estimates activity coefficients by the additive effect of contributions from all potential molecular configurations and energetic interactions between functional groups. The UNIFAC model calculates the activity coefficient of a species in a multicomponent mixture by the addition of two contributions, namely: (1) a combinatorial contribution ($ln\,\gamma_i^C$) related to molecular size and (2) a residual contribution ($ln\,\gamma_i^R$), related to the interactions between the functional groups (Narayanan, 2004, Anantpinijwatna et al., 2016) (Equation 2.32). Entropic effects caused by molecular interactions such as Van der Waals forces, surface area, and volumes are defined in the parameter $ln\,\gamma_i^C$. The $ln\,\gamma_i^R$ interaction energy of functional groups in a mixture also known as the binary interaction parameter and can be determined experimentally or computationally.

$$ln\,\gamma_i = ln\,\gamma_i^R + ln\,\gamma_i^C \qquad \textit{Equation 2.32}$$

Thermodynamic properties predicted by UNIF-LL can be coupled and fitted to experimental data, such as product concentration over time. These kinetic parameters can be formulated upon a thermodynamic framework (Straathof et al., 1992). Production of primary alcohol can be explained by rate equations while accounting for the partitioning behaviour of the substrate. Variations in partitioning coefficients and effective concentrations between the substrates of differing carbon chain lengths could explain the differences in reaction rates (Nakanishi and Matsuno, 1986).

2.5.5 Interpretation of CYP153 whole cell longer chain alkane primary hydroxylation kinetics in a 2LPS using thermodynamic modelling

Changing parameters such as the nature of the solvent, initial substrate concentration, alkane substrate chain length, alkane substrate type, and phase ratio results in changes in the substrate solubility, substrate-solvent interaction, and the effective concentration, i.e., thermodynamic activity, available to the enzyme. The solvent in a 2LPS can impact reaction kinetics by (1) a change in solvation and solubility of the substrate (solvation effects) which alters substrate availability, effective concentration, thermodynamic activity, and substrate-solvent interaction, or (2) a change in biocatalyst properties, biocatalyst deactivation, and instability and a change of intrinsic enzyme kinetics (Grosch et al., 2017). Similarly, a change in the substrate imparts the same effects because these also possess unique physicochemical properties and, thereby, unique interactions with surrounding co-solvent, aqueous molecules, or biocatalysts. It is helpful to decouple (1) from (2) by calculating or predicting the solvation effects and effective concentrations via physicochemical properties, partitioning, or thermodynamic models (Ferrario et al., 2018).

Solvation effects should be separated and discounted from biocatalyst properties and intrinsic biocatalyst behaviour. By accounting for thermodynamic phenomena, intrinsic enzyme kinetics can be distinguished from apparent experimentally measured enzyme kinetics. The latter includes limitations, such as biocatalyst performance, intrinsic enzyme kinetics, limited substrate partitioning into the aqueous phase, and biocatalyst deactivation. Therefore, experimental apparent kinetic data should be treated so that parameters determined for enzyme-substrate (ES) interaction are not skewed by substrate-solvent interaction to ensure accurate determination of intrinsic enzyme kinetics.

The experimentally determined Michaelis-Menten (MM) kinetic constant based on substrate concentration (K_m^c), rather than thermodynamic activity, depict interactions between substrate-solvent and ES. Intrinsic kinetics are complex and challenging to predict, but simpler physicochemical processes can be reliably predicted using thermodynamic models such as UNIFAC (Halling, 1994, Grosch et al., 2017). Previous studies have described and elucidated ES interactions, enzyme stability, and solvent effects on biocatalytic reactions using molecular modelling and activity-coefficient thermodynamic-based modelling approaches (Braiuca et al., 2006, Wescott and Klibanov, 1993, Grosch et al., 2017, Ferrario et al., 2018, Wangler et al., 2018). Grosch et al. (2017) identified that the solvent methyl tert-butyl ether increased the effective concentration of substrate available to the enzyme, while also effecting the enzymes performance negatively by using thermodynamic activity based kinetic modelling approach.

Thermodynamic activity coefficients can be used to calculate activity-based MM constants (K_m^a) to obtain a deeper interpretation of experimental biocatalytic 2LPS data and obtain true ES affinity. Thermodynamic-activity derived K_m^a values are 'solvent-independent' and invaluable in elucidating true ES interactions. Substrates with different solubilities, nature, or shape can be compared in relation to their interactions with the enzyme's active site, providing insight into microscopic molecular mechanisms (Ferrario and Pleiss, 2019). Ferrario et al. (2018) compared K_m^a values of different substrates as well as their respective K_m^c values. The K_m^c of P450$_{BM-3}$ on *n*-octane was the lowest of all substrates compared to aniline and oct-1-yne. The K_m^a of *n*-octane, however, was 500-fold higher than that of aniline. This indicated that the binding of P450$_{BM-3}$ and *n*-octane was influenced more by *n*-octane's poor water solubility and unfavourable interactions between aqueous molecules and *n*-octane rather than ES complementarity. In contrast, the binding between aniline and P450$_{BM-3}$ was driven by true shape complementarity.

In 2LPSs, substrate concentration defined for a kinetic model equation is confusing because the substrate is present in the organic phase, the aqueous phase, and the liquid biphasic interface. This confusion can be resolved using thermodynamic activity because the activity is identical in each phase, and the whole system is at equilibrium. Moreover, mathematical calculations and thermodynamic models such as the UNIF-LL can be used to determine the alkane concentration available to the whole cell biocatalyst in the aqueous phase. However, thermodynamic analysis of biocatalytic 2LPSs is considered an advanced topic, and additional complexity arises from inconsistent terminology among studies from different fields of biology and chemistry.

2.6 Concluding remarks

Biocatalytic 2LPSs show promising potential for highly selective (≥ 95%) primary hydroxylation of medium-chain alkanes, especially in the context of environmental sustainability. The industrial application of biocatalysis is hampered by techno-economic feasibility in large-scale practices. Whole cells have proven less expensive than purified enzymes, but this also comes with challenges. Whole cells are sensitive to harsh reaction conditions and require an aqueous environment. Compounds of poor aqueous solubility are often used as substrate feed for biocatalysis. Poor aqueous substrate solubility necessitates the application of unconventional reaction media such as the 2LPS comprising of a whole cell aqueous residing phase and a substrate residing organic phase. The poorly soluble substrate transfers across the interfacial organic-aqueous phase boundary.

The amount of substrate availability to the whole cell biocatalyst has been shown to limit reaction rates and yield, which can be investigated using kinetic and thermodynamic analysis. Aqueous phase

substrate concentration and substrate transfer across the cellular membrane are critical factors in hampering the extent and rate of the reaction. Further, an efficient reaction needs the substrate to have shape complementarity with the enzyme and efficient electron coupling, both of which contribute to intrinsic enzyme kinetics. Enzymes should be expressed at sufficient concentrations to enable adequate contact between enzyme and substrate in the whole cell cytoplasm.

CYP153A6 is an enzyme within a multicomponent system comprising electron transfer partner proteins and is widely known for its highly selective alkane hydroxylation abilities. Multicomponent enzyme systems, however, suffer from poor electron coupling. Therefore, the CYP153A13 enzyme was included to observe whether fusion protein offered superior performance in the hydroxylation of medium-chain alkanes compared to the multi-component system in which CYP153A6 functions. The major focus of the study was to investigate the impact of thermodynamic impacts on the kinetics of whole cell biocatalytic 2LPSs. In addition, the inclusion of the self-sufficient CYP153A13 fusion protein alongside the multicomponent CYP153A6 enzyme system was added to explore the potential of fusion proteins and to further bridge the knowledge gap between biology and engineering components of biocatalysis systems. This allowed a deeper understanding into the potential biological challenges that occur such as the comparison of single component and multicomponent enzyme systems.

Beyond intrinsic enzyme kinetics are parameters contributing to apparent kinetics regulated by many interdependent parameters. Investigation of these complex systems requires both biological and physicochemical considerations. Whole cell biocatalytic experimental work involves the application of microbiology skills, whereas thermodynamics requires integrating physicochemical knowledge with mathematical modelling and process simulations. Matters are further complicated because there is a prominent use of different terms defined differently across biology, chemical engineering, and mathematics fields. Moreover, studies are difficult to reproduce or compare because of a lack of standardising the reporting of biocatalytic data (Gardossi et al., 2010, Kozuch and Martin, 2012). The c*Change Centre of Excellence in Catalysis Research have developed reporting standards, which is applied throughout the study (Appendix 7.6).

Substrate and product concentrations surrounding the whole cell biocatalyst in the aqueous phase influence the phase equilibrium and substate concentration available to the whole cell biocatalyst. In turn, could affect the enzyme kinetics since the rate of the reaction may be dependent on substrate concentration. Therefore, the thermodynamics of the biocatalytic 2LPS was evaluated in terms of the phase equilibrium, whereas the hydroxylation reaction catalysed by the CYP enzymes were kinetically examined. It was attempted to decipher the impact of thermodynamically regulated phase equilibrium and solvation effects on the alkane biohydroxylation reaction kinetics and mechanisms. Additionally, substrate-solvent interactions can be accounted for in kinetic models by using thermodynamic calculation. The available concentrations of a chemical species are revealed by partitioning coefficients and solubility, which can be studied theoretically using UNIFAC modelling and computerised prediction. Several studies have sought to improve intrinsic enzyme kinetics by enzyme engineering. Other approaches include improving reaction rates by optimising environmental process conditions concerning medium composition, solvent tolerance, or oxygen transfer in large-scale reactors. However, little focus has been paid to studying the kinetics of whole cell biohydroxylation of linear alkanes in a 2LPS, especially concerning thermodynamic phenomena and substrate solubility.

2.7 Defining the Research Project

2.7.1 Project Scope

An approach was developed to investigate the kinetics of a whole cell biocatalytic hydroxylation 2LPS, focused on the conversion of alkanes to primary alcohols, coupled with thermodynamic calculations and perspective. The latter enables insight into the impact the solvent molecules may have on the binding between the substrate and enzyme, and therefore true or intrinsic enzyme properties such as enzyme-substrate affinity can be determined. The objective was to explore substrate specificity and impact on reaction kinetics of CYP153A6 and CYP153A13 towards alkanes C_8, C_9, and C_{10} because chain length influences the end-product properties, thereby implicating its value and application. CYP153A6 and CYP153A13 have shown suitability as biocatalysts for the primary hydroxylation of alkanes. CYP153A6 is highly selective to produce primary alcohols and is expressed as part of a multicomponent enzyme system. CYP153A13 has also shown high selectivity but is expressed as one component enzyme fusion, supposedly enhancing performance by improvement of electron coupling efficiencies compared to multicomponent enzyme systems. This enabled a deeper understanding of the selection of enzyme systems, albeit the major focus of the study was on investigating the impact of thermodynamic phenomena on the kinetics of whole cell biocatalytic 2LPSs.

The decreasing aqueous solubility with increasing chain length of alkanes impedes its transport across the liquid-liquid interface and cellular membrane, limiting its access to the enzyme's active site. The binding of the substrate to the enzyme to form the enzyme-substrate (ES) complex depends on the equilibrium between the ES complex and free substrate in the surrounding aqueous cytoplasmic solvent. The interactions between enzyme and substrate and the molecular interactions between the substrate and solvent molecules impact the biocatalytic performance and turnover frequency.

Aqueous substrate solubility and properties determine whether partitioning into the aqueous phase and, thus, molecular interactions between substrate and solvent are favourable for an efficient reaction. When comparing the Michaelis-Menten (MM) constants of different substrates for an enzyme, which typically indicates the affinity between enzyme and substrate, it is important to account for the thermodynamic effect of the substrate interaction with solvent. Substrate-solvent interactions and ES interactions contribute to the K_m^c constant determined experimentally for an ES pair. By comparing two K_m^c constants, it is typically assumed that the substrate with the lower K_m^c constant indicates a higher affinity for the enzyme. By calculating thermodynamic activity-based MM constants (K_m^a), solvent-substrate interactions are excluded which may reveal that the substrate does not have a higher affinity for the enzyme, but rather unfavourable interactions with surrounding aqueous cytoplasmic solvent molecules.

Therefore, enzyme reaction kinetics were interpreted in the context of thermodynamic calculations to clarify true ES affinity. To reveal true ES affinity, apparent K_m^c values were used to determine thermodynamic activity-based constants K_m^a using thermodynamic activity coefficients. Thermodynamic activity coefficients were determined for different even and odd-numbered alkanes (C_8, C_9 and C_{10}) using the UNIFAC liquid-liquid (UNIF-LL) thermodynamic model in process flowsheet simulations. Insights into substrate availability and partitioning into the aqueous phase containing the whole cell biocatalyst were also provided by the UNIF-LL process flowsheet models, which helped to determine whether the reaction was limited by or independent of substrate concentration.

From an industrial perspective, one-pot processes are becoming increasingly attractive to decrease the time, resources, and solvents needed to achieve techno-economic feasibility and improve the

environmental sustainability of fine chemical synthesis. Therefore, research into the thermodynamics and kinetics of systems that include multiple enzymes and substrates is essential. The investigations sought to assist in understanding true enzyme kinetics in a 2LPS. This is relevant in one-pot processes because reactions of the ES pair that has the lowest K_m is favoured. Methods were investigated that reveal true K_m constants and intrinsic enzyme kinetics, expanding on classical water-based enzyme kinetics to whole cell biocatalytic primary hydroxylation of medium chain alkanes in a 2LPS. Further, the calculation-based approaches were integrated with experimental whole cell 2LPS studies to characterise and compare the performance of CYP156A6 and CYP153A13 enzyme systems across alkane chain lengths from C_8 to C_{10}.

Experimental whole cell studies involved determining the product concentration accumulated at different time points of the biocatalytic reaction for each ES pair, along with the determination of enzyme concentration. This enabled the determination of turnover frequency and therefore preliminary understanding of the substrate scope of each enzyme. Product as a function of time graphs were fitted to the MM model to determine experimentally the K_m^c constants and maximum reaction velocities. Also, CYP153A13 is a fusion enzyme which is known for its susceptibility to undesired proteolytic cleavage, and therefore it was purified and studied using protein viewing gel tools to characterise its expression.

2.7.2 Problem Statement and objectives

Several CYP153-based whole cell biocatalytic hydroxylation studies have focused on improving intrinsic enzyme kinetics, biological factors such as improved expression or electron coupling, mass transfer into and out of the cell, and environmental conditions. The effect of thermodynamically governed substrate partitioning into the aqueous phase and molecular interaction between the substrate-solvent on the enzyme kinetics of whole cell medium chain length alkane biohydroxylation reactions has not been extensively studied. Thermodynamic and physicochemical phenomena related to substrate-solvent interactions, substrate supply to the whole cells, substrate solubility, substrate type, and chain length were investigated. CYP153A6 and CYP153A13 were selected because these enzymes are well known to be suitable for primary biohydroxylation of medium chain length alkanes.

Furthermore, multicomponent systems such as CYP153A6, although highly selective for the activation of the primary carbon of alkanes, are susceptible to poor coupling efficiencies, resulting in undesired by-products formation such as H_2O_2, deactivating CYP153 enzymes and consequently decreasing turnover. Therefore, the potential gain in turnover by using the engineered fusion CYP153A13 was investigated and compared to the highly selective CYP153A6. The expression of CYP153A13 was also characterised to determine whether it had susceptibility to protein truncation.

The alkane substrate scope of CYP153A6 and CYP153A13 was evaluated. The evaluation was expanded by investigating the effect of the solubility, chain length, and type of substrate on turnover frequency as well as decoupling substrate-solvent interactions from enzyme-substrate (ES) interactions to reveal true ES affinity, informing the kinetics of C_8 C_9 and C_{10} alkane biohydroxylation in 2LPSs. UNIF-LL process flowsheet models also provided information on substrate availability, which could be used to determine whether reactions were limited by substrate concentration or other factors such as enzyme expression or ES affinity. Constraints from both biological and chemical engineering fields were identified, and interdisciplinarity was improved by integrating knowledge from these fields.

Apparent kinetic MM constants (K_m^c) and rates (turnover frequency or maximum reaction velocity) take into account several interactions, including ES, substrate-solvent, and enzyme-solvent. These were determined by analysing the experimentally determined concentration of accumulated product as a

function of time and fitting the profiles on these graphs to the MM model using nonlinear regression. From this, the K_m^c constants were determined, and maximum reaction velocities could be compared amongst the ES pairs studied.

Few studies have explored substrate-solvent interaction for CYP153 whole cell biocatalysis on medium-chain alkanes. Substrate-solvent interactions were decoupled from enzyme-substrate interactions by using thermodynamic activities to determine true intrinsic K_m^a constants. The objective was to determine the intrinsic kinetic properties of CYP153A6 and CYP153A13 biocatalysing the primary hydroxylation of *n*-octane, *n*-nonane, and *n*-decane, respectively. This facilitated understanding of the impact of different alkane chain lengths, substrate type, and their respective solubilities on biocatalytic performance on enzyme-substrate and substrate-solvent levels. Overall, the lack of knowledge in the field of solvation effects and its impact on the reaction kinetics in whole cell biocatalytic 2LPSs was addressed. The following steps were undertaken to address the project objectives:

- Genetically engineer *E. coli* host with CYP153A13 gene to prepare whole cell biocatalyst.
- Test the success of recombinant expression of CYP153A13 in *E. coli* by characterising protein concentration and functionality using CO-difference spectral assays and SDS-PAGE, respectively.
- Perform preliminary substrate scope analysis to determine the functional CYP153A6 and CYP153A13 whole cell biocatalytic activity on *n*-octane, *n*-nonane, and *n*-decane.
- Determine apparent kinetic constants (K_m^c) of each enzyme-substrate pair by fitting the product concentration as a function of time curve to the Michaelis-Menten equation using least squares nonlinear regression (NLR).
- Determine thermodynamic activity-based kinetic parameters (K_m^a) of each enzyme-substrate pair, which excludes the effects of solvent-substrate interactions on enzyme-substrate interactions, to reveal intrinsic CYP153A6 and CYP153A13 kinetics.
- Computationally predict substrate mole fractions in aqueous phase using UNIF-LL process flowsheet models to determine partitioning coefficients and thereby determine the substrate concentration available to the whole cell biocatalyst. This enabled insights into whether a reaction is limited by substrate concentration or not.

2.7.3 Research hypothesis and key questions

Whole cell biocatalytic 2LPSs suffer from low volumetric rates and reaction rates due to several interdependent factors. Identifying which are the most limiting is essential to pave the way for larger-scale and more efficient biocatalytic processes. It was hypothesised that substrate supply and access to enzymes is one of the most constraining factors on biocatalytic performance due to limited transport across the cell membrane and liquid-liquid interface, especially when substrates have poor aqueous solubility. As the chain length of alkanes increases, solubility decreases; therefore, longer chain alkanes are expected to show lower biocatalytic volumetric rates and reaction rates, regardless of favourable enzyme-substrate affinity.

Molecular interactions between substrate and solvent impact experimentally determined apparent K_m^c constants. The UNIF-LL process flowsheet model was used to determine activity coefficients and, in turn, thermodynamic activity-based Michaelis-Menten K_m^a constants of each ES pair to determine true affinity between the enzyme and substrate. Therefore, kinetic data for the biocatalyic hydroxylation of alkanes were explored in relation to thermodynamic activity rather than classical enzyme kinetic studies

which are based on monophasic and ideal dilute aqueous conditions using purified enzyme systems. Moreover, the UNIF-LL process flowsheet model was used to examine the substrate supply to the whole cell biocatalyst across the two immiscible liquid phases to determine whether the reactions were limited by substrate availability.

2.7.3.1 Purpose & Function

CYP153A13 has shown promising activity and high affinity to medium-chain alkanes and was expressed as a single polypeptide unit since fusion enzymes have shown superior turnover frequencies due to improved electron coupling efficiencies. In fusion CYPs, electron transfer proteins are in closer proximity than in multicomponent enzyme systems. CYP153A13 was compared to the CYP153A6 enzyme system, a well-studied multicomponent enzyme system known for its high selectivity for the primary hydroxylation activity on medium-chain alkanes. Therefore, the superiority of fusion enzymes was investigated using molecular biology techniques and a comparison of TOFs over time.

Enzyme-substrate affinities were investigated using a range of chain-length alkanes in a 2LPS using whole cell biocatalysts. Since linear alkanes pose solubility issues, it is easy to assume that reactions are limited by substrate availability and solubility. In the present study, it was examined whether the reactions were, in fact, limited by substrate availability and if not, what other factors might be limiting. Alkane chain length and nature (odd versus even) were alternated to facilitate the understanding of substrate solubility on biocatalytic performance. A method using a thermodynamic model and calculations was developed to determine the most suited substrates for the two enzymes under investigation by considering substrate availability as well as enzyme-substrate affinity.

2.7.3.2 Key Research Questions

1. Does the fusion single component CYP153A13 enzyme show superior biocatalytic qualities such as turnover frequency (TOF) compared to the well-known multicomponent CYP153A6 system?
2. How do alkane chain length and solubility contribute to the resultant turnover frequencies and reaction velocities in 2LPSs?
3. Can least squares non-linear regression be used to fit product concentration as a function of time data to the Michaelis-Menten model to accurately determine apparent kinetic constants and maximum reaction velocities of whole cell alkane biohydroxylation?
4. How do the apparent K_m^c constants differ from true K_m^a constants for each enzyme-substrate pair differ in terms of enzyme-substrate affinity?
5. What is the influence of substrate solubility and availability on the turnover frequency of the respective reactions?
6. How do the partitioning coefficients of the alkane's differ as chain length increases or the nature of the alkane changes i.e., odd versus even chain lengths?
7. Are the reactions of each enzyme-substrate pair limited by substrate availability and concentration?

3 Approach to Project and Methodology

The research methodology included wet-laboratory experiments, computer-based process simulations, and modelling. Wet-laboratory-based experiments included the preparation of recombinant whole cell biocatalysts using genetic manipulation, inoculation, selection and growth, induction of protein expression, protein analyses, and alkane biotransformation tests. Product concentration as a function of time data was processed to estimate the apparent kinetic parameters, maximum reaction velocity (v_{max}) and Michaelis-Menten (MM) constant (K_m^c) by using least squares non-linear regression (NLR). UNIF-LL process simulations were used to determine the impact of parameters such as solvation effects, solubility, solvent selection, and temperature on aqueous phase substrate availability in the aqueous phase. UNIF-LL modelling was used to determine the intrinsic thermodynamic-activity-based MM constants K_m^a as well as the impact that aqueous substrate concentration according to its partitioning behaviour between the two immiscible phases has on the kinetics of the reaction. K_m^a values were used to decouple substrate-solvent effects from the enzyme-substrate (ES) interactions to attain intrinsic and true ES compatibility of the biocatalytic hydroxylation 2-Liquid-Phase-System (2LPS).

3.1 Experimental materials and methods

All chemicals and antibiotics were obtained from Sigma-Aldrich, Roche, or Merck (South Africa).

3.1.1 Microorganism selection and growth

Various experimental methods were performed to investigate and compare CYP153A13's expression and functionality to the CYP153A6 and CYP102A1 F87V-A328F-R966D-W1046S mutant (kindly provided Prof M.S Smit's Biocatalysis Laboratory, University of Free State, South Africa). For *in vivo* biocatalytic studies, CYP153A6 and CYP153A13's expression was induced and the whole cell biotransformation activity investigated. For *in vitro* biocatalytic studies, CYP153A13 and CYP102A1 mutant were purified, and their biotransformation activity studied.

Recombinant organisms were prepared by transforming BL21 (DE3) *Escherichia coli* host cells with plasmid constructs (Figure 3.1). The CYP153A13 in pET28b (+), was transformed into *Escherichia coli* BL21 (DE3) using heat shock transformation (Appendix 7.1). CYP153A6 and CYP102A1 mutant whole cell biocatalysts had previously been prepared in the University of Cape Town's laboratory using pET28b(+) plasmid constructs.

The CYP153A6 plasmid construct comprised the complete operon from *Mycobacterium* sp. HXN-1500 which consists of the CYP153A6, ferredoxin reductase (FdR), and ferredoxin (Fdx) genes (Olaofe et al., 2013).The CYP153A13 plasmid construct used containing the CYP153A13a gene isolated from *Alcanivorax borkumensis* and the P450 reductase domain (RhFred) gene isolated from *Rhodococcus* sp. NCIMB 9784 (Bordeaux et al., 2011) in a pET28b(+) vector (Figure 3.1). CYP153A6 and CYP153A13 plasmid constructs both containing kanamycin resistance gene are shown in Figure 3.1. Stock cultures of both whole cell biocatalysts CYP153A6 and CYP153A13 were maintained in glycerol (25%) at -60˚C.

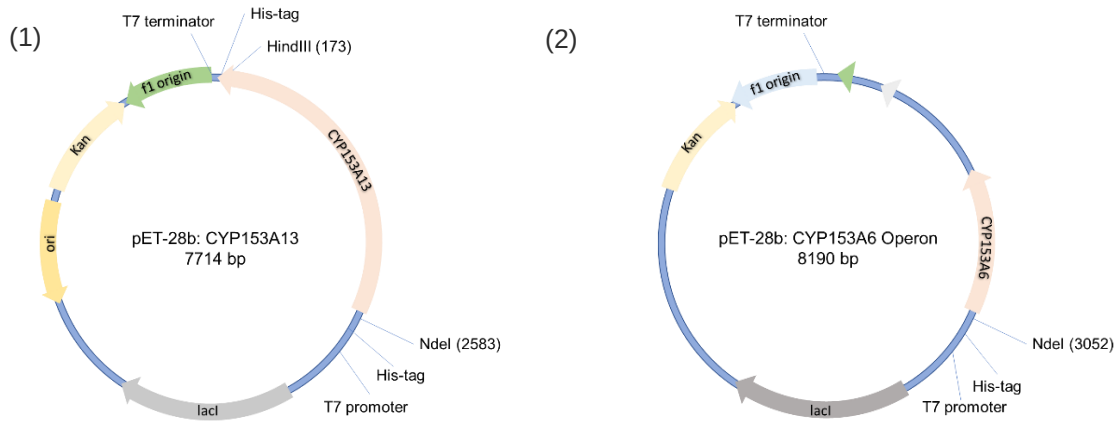


Figure 3.1. Plasmid maps of pET-28b consisting of (1) Single-component CYP153A13a gene isolated from *Alcanivorax borkumensis* fused to the P450 reductase RhFred domain and (2) Multicomponent CYP153A6 gene along with the ferredoxin and ferredoxin reductase genes (Olaofe et al., 2013)

3.1.2 Pre-inoculum cultures

Luria Bertani (LB) broth was prepared by adding 10 $g \cdot L^{-1}$ tryptone, 5 $g \cdot L^{-1}$ NaCl, and 5 $g \cdot L^{-1}$ yeast extract in distilled water. An autoclave then sterilised the prepared LB medium for 20 minutes at 121˚C. A 10 mL volume of pre-sterilised LB media was placed in McCartney bottles, and pre-filtered kanamycin stock was added to obtain a final concentration of 30 $\mu g \cdot mL^{-1}$ kanamycin to maintain selective pressure for positive recombinant transformants. Both whole cell biocatalysts were added to the media from previously maintained glycerol stocks. Pre-inoculum cultures were incubated in McCartney bottles for 12 hours at 37˚C and 120 rpm.

3.1.3 Main inoculum and induction of protein expression

A total volume of 90 mL of pre-sterilised LB medium was added to pre-sterilised 500 mL Erlenmeyer flasks. Pre-filtered kanamycin stock culture was added to the medium to obtain a final kanamycin concentration of 30 $\mu g \cdot mL^{-1}$ to maintain selective pressure of positive transformants. The 10 mL pre-inoculum culture was added to the 90 mL LB media volume in the Erlenmeyer flask to obtain a 10% v/v in a working volume of 100 mL in a 500 mL flask.

Once the mid-exponential phase was reached, indicated by an OD_{600} of 0.8 to 1, the culture was supplemented with 0.5 mM 5-aminolevulinic acid (5-ALA), 50 µM $FeCl_3.6H_2O$ and protein expression was induced by adding 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG) (Gudiminchi et al., 2012). Once protein expression was induced, cultures were incubated at 20˚C and 200 rpm for 24 hours for CYP153A6 and CYP153A13 whole cell biotransformation studies (See Chapter 3.1.6.2).

3.1.4 Purified enzyme characterisation and purification

CYP153A13 was purified alongside CYP102A1 mutant (a well-known soluble fusion enzyme) for SDS-PAGE gel, heme stain gel, and CO-difference spectral analyses. *In vitro* activity and expression of crude lysate and purified (desalted and dialysed) CYP153A13 and CYP102A1 mutant were investigated (Chapter 3.5.1). For pre-inoculum cultures, single colonies containing the CYP153A13 and CYP102A12 mutant plasmid constructs were respectively selected and grown for 12 hours at 37˚C and 160 rpm in a volume of 50 ML pre-sterilised LB media containing pre-filtered 30 $\mu g \cdot mL^{-1}$ kanamycin to maintain selective pressure for positive recombinant transformants. Pre-inoculum cultures were added to 500 mL pre-sterilised ZYP-5052 auto-induction medium containing 1 mM 5-ALA and 0.05 mM $FeCl_3$, 30 $\mu g \cdot mL^{-1}$ kanamycin and the cultures incubated for 48 hours at 20˚C at 200 rpm.

All samples were kept on ice throughput the procedure to avoid proteolysis. Cell pellet was harvested by centrifugation at 7000 rpm at 4˚C for 10 minutes. The cell pellet resuspended to obtain a concentration of 1 g of cell wet weight per 10 mL, in 200 mM Tris-HCL lysis buffer at pH 7.4 containing protease inhibitors (0.7 $\mu g \cdot mL^{-1}$ pepstatin, 0.1 $\mu g \cdot mL^{-1}$ perfabloc SC, 0.5 $\mu g \cdot mL^{-1}$ leupeptin). The resuspended cell pellet was lysed by sonication (misonix sonicator 3000) on ice for 8 minutes (30 seconds on and 30 seconds off for cooling). The soluble protein fraction was separated by centrifugation at 17 000 rpm (20 000 g) for one hour using a JA-20 centrifuge. The supernatant was filtered through a 0.45 µM filter to form a cleared lysate. The remainder of the supernatants were used as crude extracts or to further purify proteins.

3.1.4.1 Nickel affinity purification

The samples obtained from the previous step were subjected to a IMAC chromatography purification using an AKTA explorer purification system (Amersham pharmacia biotech), a Frac-900 collector and a 5 mL HiTrap™ Chelating HP column (GE Healthcare, Sweden) loaded with Ni^{2+} ions. Before the CL was loaded, the column was equilibrated with 5x column volumes of 50 mM Tris-HCL, 0.5 M NaCl, 20 mM imidazole (pH 7.4) equilibration buffer. The protein bound to the column was eluted by applying a linear gradient of 0 to 500 mM imidazole in 50 mM Tris-HCl buffer at a pH of 7.4 to collect protein fractions. The fractions that contained the highest concentration of protein were desalted on a HiPrep™ 26/10 desalting column (GE Healthcare) or by dialysis overnight in 200 mM Tris HCl buffer at pH 7.4. An ultrafiltration step was used to further concentrate the purified protein samples in an Amicon Ultra-15 Centrifugal Filter unit. Each protein sample (clear lysates, desalted, dialysed) was examined for activity and purity using biotransformation studies and SDS-PAGE gels, respectively. Heme stain gels were used to confirm the presence of the CYP active site (Thomas et al., 1976, Francis Jr and Becker, 1984). Functional protein concentrations were determined using CO-difference spectra using an extinction coefficient of 91 $mM^{-1}cm^{-1}$ for the 450 minus 490 nm peak (See Chapter 3.2.2).

3.1.4.2 SDS-PAGE Coomassie stain gel

Purified proteins were viewed by carrying out sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) using the Mini-PROTEAN® 3-system (Bio-Rad). The separating gel consisted of 10% resolving gel and a 4% stacking gel. To make sample buffer, a volume of 50 μL of 2-Mercapto-ethanol was added to a volume of 950 μL of 200 mM Tris-HCL buffer at pH 7.4. A volume of 10 μL of each sample (containing 2 mL of sample buffer and 8 mL of purified protein) was loaded into the wells of the gel. The gel was run under 150 V for one hour in 25 mM Tris, 192 mM glycine and 0.1% (w/v) SDS running buffer at pH 8.3. Coomassie® brilliant blue R-250 was used to stain the gel. In lane 1, PageRuler™ pre-stained Protein Ladder (Thermoscientific) was used as molecular weight marker from 10 to 180 kDa.

3.1.4.3 Heme stain gel

The gels were stained for the detection of peroxidase activity to confirm the presence of the heme group in the active site of the CYP enzymes, as described by Thomas et al. (1976). A mass of 30 mg of *N, N, N', N'*- tetramethylbenzidine (TMB) (6.3 mM) was dissolved in 15 mL methanol and mixed with a volume of 35 mL of 0.5 M Sodium Acetate (pH 5). The gel was washed with distilled water for five minutes and placed carefully in the stain. The gel was left in the stain for 30 minutes in the dark with gentle shaking. A volume of 300 µL of hydrogen peroxide to make a final concentration of 30 mM was added to the stain. After five to thirty minutes, the blue bands on the gel appeared and were observed (Thomas et al., 1976). Non-denaturing conditions were used to ensure successful staining.

3.1.5 Biotransformation conditions

3.1.5.1 In vitro activity

Reactions were carried out at 20˚C at 200 rpm in 60 mL amber vials containing 300 μL of either cleared lysate, desalted or dialysed samples, along with 100 mM glucose, 5 U glucose dehydrogenase· mL^{-1}, 1 mM NADPH and 200 mM Tris-HCl at pH 7.4. In vitro activity of CYP153A13 and CYP102A1 crude and purified extracts were tested by placing 300 µL of either crude, desalted or dialysed samples in amber vials respectively. Thereafter, a volume of 200 µL octane and 50 µL BEHP were added to initiate the reaction.

3.1.5.2 In vivo activity using Low Cell Density Cultures

Sodium phosphate buffer at a concentration of 200 mM at pH 7.2 was prepared, containing 100 mM glucose and 0.8% glycerol. The buffer was sterilised using an autoclave for 20 minutes at 121˚C. After protein expression was induced for 24 hours, *E. coli* BL21(DE3) cells were harvested by centrifugation at 7000 *g* at 4˚C for 10 minutes. The cell pellet was resuspended in the sodium phosphate buffer at a concentration of 3-5 $g_{DCW} \cdot L_{BRM}^{-1}$. The resuspended cell pellet in phosphate buffer is called the Bioreaction Mixture (BRM).

The concentration of active CYP in the BRM samples was analysed further (Chapter 3.2.2). Multiple one millilitre aliquots of the BRM were placed into pre-labelled 60 mL sterile amber screw cap vials. The substrate of either *n*-octane, *n*-nonane, or *n*-decane was pre-mixed with carrier solvent BEHP in a volumetric ratio of 2:1, respectively, which made up the 'organic phase'. A volume of 300 µL organic phase aliquots was placed into the amber vials containing BRMs of respective whole cell biocatalysts. The screw caps were tightened onto the amber vials and placed on an orbital shaker at 200 rpm and 20˚C. These were standard biotransformation conditions unless stated otherwise. Vials were sacrificed at specified time points (3, 6, 9, 12, 24, 48, and 72 hours) for further chemical and analytical analyses. Biological repeat one (BR1) included sampling for a total reaction time of 3 to 72 hours. Two independent biological repeats (biological repeat 2 and 3) were performed in duplicate but included sampling points from 24 to 72 hours and 3 to 12 hours, respectively.
To ensure that all product formation was due to enzymatic catalysis, a separate biotransformation experiment was performed that included negative controls: uninduced CYP and empty vector cultures.

3.1.5.3 In vivo activity using High Cell Density Cultures

The same method was followed as 'Low Cell Density Cultures' preparation, except the concentration of cells prepared for the BRM was increased to 78-87 $g_{DCW} \cdot L_{BRM}^{-1}$. High cell density cultures were performed at 24- and 48-hour sampling time points.

3.2 Analytical experimental methods

3.2.1 Quantification of cell growth and biomass

3.2.1.1 Cell turbidity

The cellular growth of inoculum cultures was monitored by measuring cell culture turbidity at a wavelength of 600 nm using a Thermoscientific Genesys 10S UV-Vis spectrophotometer. Higher biomass concentration samples taken during the late-exponential phase (OD_{600} of 0.8 to 1) were diluted to below an absorbance reading of 0.8 to prevent inaccuracies explained by the Beer-Lambert law. The original absorbance of the diluted cell culture sample was calculated by obtaining the product of the measured absorbance value of the diluted sample by the dilution factor.

3.2.1.2 Dry cell weight

The preparation of the bioreaction mixture (BRM) required the suspension of a specific cell concentration to obtain a certain mass of dry cell weight (DCW) per volumetric unit of the sodium phosphate buffer. Cell turbidity can be measured quickly (Chapter 3.2.1.1), while DCW takes at least 48 hours. Therefore, a standard curve was constructed to establish the relationship between dcw and OD_{600}, and it was shown that the method showed an error of 5 and 4%, respectively (Appendix 7.2). The OD_{600} of samples was measured quickly (c_1) to determine the correlated DCW by using the standard curve. The final desired concentration (c_2) for BRM in DCW measured in $g_{DCW} \cdot L_{BRM}^{-1}$ was prepared according to the initial cell concentration of the recombinant (c_1) (Equation 3.1).

$$c_1 v_1 = c_2 v_2 \qquad \textit{Equation 3.1}$$

3.2.2 Quantification of CYPs

The concentration of functional CYP proteins was determined using the CO-difference spectral assay described by Guengerich et al. (2009). This technique was applied to purified enzyme and whole cell samples. Carbon monoxide (CO) gas tightly binds with only the ferrous (Fe^{2+}) form of the heme active site of cytochrome P450 (CYP) enzymes, forming a complex that produces a maximum absorbance at a wavelength of 450 nm. Hence, properly folded and active CYP protein are characterised by a peak at a wavelength of 450 nm due to a cysteine thiolate axial ligand coordinated to the heme iron and this is proportional to the concentration of active enzyme. Inactive or poorly folded CYPs show a lower or no peak at the 450 nm wavelength and an increased peak at 420 nm. Therefore, the CYP CO-difference spectral assay accurately represents the enzyme's integrity, expression, activity, and functionality.

Before biotransformation, 1 mL of BRM samples were split into two test tubes (Guengerich et al., 2009). The contents of one of these tubes were sparged with CO gas for 30 minutes (sample tube), and the other reference tube had no CO treatment. The reference and sample tubes were reduced artificially by adding 1 mg of sodium dithionite. After that, the reference tube contains the ferrous form of reduced CYP, while the sample tube contains the same ferrous CYP bound to CO.

The sample and reference tubes were scanned to obtain a 400-500 nm spectrum with a UV-Vis spectrophotometer. The concentration of active CYP in the whole cell samples was calculated by obtaining the difference between the scanned spectrums of samples treated with CO and not treated with CO at the previously determined extinction coefficient $(91\ mM^{-1}cm^{-1})$(Equation 3.2)

$$[CYP] = \frac{\Delta A450 - \Delta A490}{0.091} \qquad \textit{Equation 3.2}$$

The CO-difference spectral analyses of some biological repeats in this study were unsuccessful due to high biological background noise or inadequate enzyme concentrations in some of the biological repeats. Therefore, CYP concentrations were taken as nominal average ± SE 95% CI of three other biological repeats with a relative standard error of 10 to 15% (Appendix 7.3).

3.2.3 Extraction and analysis of primary alcohol

During the biotransformation reaction, amber vials were sacrificed at the specified time points, and samples were subjected to ethyl acetate (EtAc) solvent extraction. Product concentration was measured using gas chromatography (GC), which had a 15% or less average relative standard error of the mean. Samples were placed on ice in all steps during extraction to avoid product evaporation.

The contents of the amber vial were extracted using 500 µL of EtAc containing an appropriate internal standard. The contents were vortexed in Eppendorf tubes for ten minutes at room temperature, before separation into two layers (organic and aqueous) by centrifugation at 13 000 rpm for three minutes at 4˚C. The top organic phase layer was removed carefully and placed into separate sterilised Eppendorf tubes. The extraction procedure was repeated by adding another 500 µL EtAc to the aqueous phase layer to ensure that any residual primary alcohol was extracted into the organic phase. The total volume of EtAc extract, therefore, was 1 mL.

A 200 µL aliquot was removed from the total organic extract of 1 mL volume, placed into clean GC vials, and subjected to GC analysis. A Varian 3900 series gas chromatograph equipped with a flame ionisation detector and a non-polar Varian factor four column (VF-1ms) containing dimethylpolysiloxane with measurements of 15 m $\times$ 0.25 mm. The sample volume injected was 1 µL, the injection temperature was 280˚C, and a nitrogen carrier gas flowed through the column at 15 mL·min^{-1}. The temperature programme was set to the following conditions: an isotherm was maintained at 120˚C for five minutes, and the temperature was then increased by 20˚C min^{-1} to 280˚C, followed by a seven-minute isotherm (Appendix 7.4).

Primary alcohol analysed generated peaks at specific retention times in the gas chromatograms. Subsequently, primary alcohol concentrations were calculated using standard curves ranging from primary alcohol concentrations of 20 mM to 100 mM. The solvent used to dilute the primary alcohol contained an identical concentration (2 mM) of internal standard as the solvent used for sacrificial extractions of the BRM (Appendix 7.5). For analysis of 1-octanol and 1-nonanol, a 1-decanol internal standard was used and for 1-decanol analysis a 1-octanol internal standard was used The y-axis of the standard curve comprised the ratio of the peak area of the primary alcohol product to the peak area of the chosen internal standard, and the x-axis was the independent variable which was alcohol concentration. The standard curves constructed for the different alcohol products had a 3% and less average relative standard error of the mean.

3.3 Data processing of results

Data processing of product concentration as a function of time was performed to obtain process performance metrics such as specific product concentration or Turnover Number (TON), specific activity or Turnover Frequency (TOF), volumetric rate, biocatalytic efficiency, and maximum specific biocatalytic activity (*viz.* Table 2.2 in Literature Review 4.4). The TOF was interchangeably used with the catalytic rate constant k_{cat}, which is a first-order rate constant and indicates the number of molecules converted to product per unit of time when an enzyme works at maximum efficiency. The data was reported according to the standardised approach by 'The Paraffin Activation' research division of the c*Change Centre of Excellence in Catalysis Research (Appendix 7.6).

Since it was observed that there was a large amount of intercultural variability between biological repeats, experimental data was drawn from a single biological repeat at a time. This method enabled a fair comparison of whole cell biocatalytic activity and impartial conclusions of the impact of substrate solubility and alkane chain length on whole cell biocatalytic performance. Biological Repeat 1 (BR1) included whole cell biocatalysts CYP153A6 and CYP153A13 on *n*-octane and *n*-nonane at sampling points 3, 6, 9, 12, 24, 48, and 72 hours. To validate the trends observed in BR1, two additional biological repeats were included, BR2 and BR3. BR2 included CYP153A6 and CYP153A13 whole cell biocatalysts on *n*-octane, *n*-nonane, and *n*-decane at sampling points 3, 6, 9, and 12, whereas BR3 sampled at 24, 48, and 72 hours.

Enzyme-substrate (ES) pair biocatalytic performance was evaluated by comparing single-time point data collected at the 24-hour sampling point to illustrate the yield and capability of an enzyme on alkane substrates. To enable comparison to other studies, the TOF at 24-hour reaction time was reported, regardless of the maximum TOF. Data of TOF over time were analysed to gain a deeper comprehension of the reaction kinetics of ES pairs and to obtain a clear view of the reaction rates over time. Michaelis-Menten kinetic parameters were drawn from product concentration as a function of time data and processed further to exclude solvent effects from the MM constant.

3.4 Determination of apparent enzyme kinetic analyses CYP153A6 and CYP153A13 whole cell biohydroxylation on a series of medium chain alkanes

3.4.1 Determination of experimental substrate concentration-based apparent Michaelis-Menten enzyme kinetic parameters v_0, v_{max} and K_m^c

Michaelis-Menten (MM) kinetic models were investigated to determine the relationship between alkane concentration and alkane hydroxylation rate. The MM model describes the relationship between initial reaction velocity (v_o) and substrate concentration (Equation 3.3). Usually, the initial reaction velocities of multiple experiments in monophasic systems are measured, in which substrate concentrations are easily controllable. In 2LPSs, the substrate concentration is not as easily controlled, and excess substrate was applied to enable maximum substrate partitioning into the aqueous phase. Substrate concentration exposed to the catalyst was governed by phase equilibrium. Hence, the conventional route of experimental construction of the MM plot, by plotting the initial reaction velocities measured from different substrate concentrations, could not be applied.

$$\frac{dP}{dt} = v = \frac{v_{max}[S]}{K_m^c + [S]} \qquad \textit{Equation 3.3}$$

The mass balance equation shows that the reaction velocity is also equivalent to the rate of product formation, which equals the rate of substrate disappearance (Equation 3.4). Apparent kinetics assumed that all substrate is accessible to the enzyme's active site.

$$v = \frac{d[P]}{dt} = -\frac{d[S]}{dt} \qquad \textit{Equation 3.4}$$

3.4.1.1 Initial reaction velocity (v_o)

The v_o for each enzyme-substrate (ES) pair was determined at a substrate concentration governed by phase equilibrium. The tangent's gradient to the linear section of product concentration as a function of time curve to calculate v_0, since factors such as biocatalyst deactivation emerge in the later stages of a reaction (Equation 3.5).

$$v_0 = Gradient\ of\ tangent\ to\ linear\ section\ of\ [P]\ versus\ t \qquad \textit{Equation 3.5}$$

3.4.1.2 Maximum Velocity and MM constant (v_{max}, K_m^c)

To determine apparent substrate concentration-based apparent MM kinetic parameters v_{max} and K_m^c of whole cell biocatalysts on alkanes in a 2LPS, least squares nonlinear regression (NLR) was used to fit product concentration as a function of time data to the MM model. Apparent kinetic parameters were

calculated using the total excess substrate placed into the 2LPS. The approach was based on a few assumptions: (1) substrate depletion was never reached because it was applied in excess, (2) whole cell biocatalysts maintained maximum cellular health throughout the whole 72-hour reaction, with enzyme deactivation or cellular membrane disintegration from compound toxicity (3) kinetic parameters could be determined using least squares NLR at a single substrate concentration exposed to the whole cell biocatalyst which was governed by thermodynamic phase equilibrium and lastly (4) 100% of the substrate converted resulted in the desired product with no by-products or substrate loss.

Initial v_{max} estimates for NLR were calculated by the product of the enzyme concentration $[E]$ at t_0 and the catalytic rate (k_{cat}) (Equation 3.6) (Tables 4.11 and 4.12 in Chapter 4.3). The maximum TOF and k_{cat} were interchangeable, and both measured by $\mu mol_{1-alkanol} \cdot \mu mol_{P450} \cdot min^{-1}$. K_D constants of CYP153A6 for alkanes determined by Funhoff et al. (2006) were used for initial estimates for K_m^c constants. K_D constants of CYP153A13 on alkanes, determined by Bordeaux et al. (2011), were used as initial estimates.

$$v_{max} = k_{cat}[E] \qquad \textit{Equation 3.6}$$

The total initial substrate concentration ($[S]_0$) was calculated for each alkane 2LPS according to Equation 3.7. The working volume (WV) is the volume available for reaction, which includes the aqueous volume and alkane but excludes the BEHP solvent. This concentration differs from the effective concentration that is based on thermodynamic activity of the substrate, discussed below (Chapter 3.5.1). In contexts that the kinetics are determined using total substrate concentration, the term ‘apparent kinetics’ applies, whereas intrinsic kinetics are determined using thermodynamic activity.

$$[S]_0 = \frac{Total\ amount\ of\ substrate}{WV} \qquad \textit{Equation 3.7}$$

The product concentration as a function of time was transformed and fitted to Equation 3.3 by least squares NLR using Solver Add-in on Microsoft Excel (Table 3.3, Equations 3.7 to 3.11). This enabled the estimation of kinetic parameters v_{max} and K_m^c by minimising the objective function, which was the sum of squared residuals (Equation 3.11).

Table 3.3. Nonlinear regression table used on Microsoft Excel Solver Add-in to estimate kinetic parameters by fitting experimentally measured product concentration to the Michaelis-Menten output concentration optimised by minimising the objective function, which was the sum of squared residuals.

Time (hours)	Substrate concentration ($\mu mol_{alkane} \cdot L_{WV}^{-1}$) (Model)	Product concentration ($\mu mol_{1-alkanol} \cdot L_{WV}^{-1}$) (Model)	Product concentration ($\mu mol \cdot L_{WV}^{-1}$) (Experimental)	([P]model-[P]exp)^2
0	$[S]_0$ *(Equation 3.7)*	0	0	0
t_{n+1}	$[S]_{n+1}$ *(Equation 3.8)*	Model $[P]_{n+1}$ *(Equation 3.9)*	Experimental $[P]_{n+1}$	*Equation 3.10*
Objective function				*Equation 3.11*

$$[S]_{n+1} = -\left(\frac{v_{max}[S]_n}{K_m + [S]_n}\right)(t_{n+1} - t_n) + [S]_n \qquad \textit{Equation 3.8}$$

$$[P]_{n+1} = \left(\left(\frac{v_{max}[S]_n}{K_m + [S]_n}\right)/WV\right)(t_{n+1} - t_n) + [P]_n \qquad \textit{Equation 3.9}$$

$$([P]_{t,exp} - [P]_{t,model})^2 \qquad \textit{Equation 3.10}$$

$$SSR = \sum_i ([P]_{t,exp} - [P]_{t,model})^2 \qquad \textit{Equation 3.11}$$

To determine how well the data fit the model, R^2 coefficients were calculated by subtracting the ratio of the SSR to the sum of the square of the y-variable (product concentration) minus the mean of the y-variable (mean product concentration from) one (Equation 3.12) (Nagar et al., 2014).

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2} \qquad \textit{Equation 3.12}$$

3.5 UNIFAC solvation-corrected enzyme kinetic parameters of 2LPS CYP153 whole cell biohydroxylation of medium chain alkanes

3.5.1 Determination of activity coefficient using UNIFAC Liquid-Liquid Group Contribution Method

The impact of substrate-solvent interactions and substrate availability kinetics on alkane biohydroxylation in 2LPS was explored using thermodynamic calculations to determine the true enzyme-substrate (ES) affinities. The thermodynamic activity (a) or effective concentration of a compound equates to the product of the thermodynamic activity coefficient (γ) and the mole fraction (x) (Equation 3.13) (Ferrario et al., 2018, Halling, 1994). The respective activities of the alkane substrates in each liquid phase were determined by predictive simulations and theoretical calculations. For compounds of poor water solubility, γ is assumed to be independent of concentration and calculated at infinite dilution (Ferrario et al., 2018).

$$a = (\gamma)(x) \qquad \textit{Equation 3.13}$$

The a was determined using UNIF-LL modelling process simulations on Aspen Plus (V12). The experimental 2LPSs were simulated using process flowsheets designed to mimic laboratory experiments without whole cell biocatalyst, and the activity coefficients of alkanes were determined (Figure 3.2). Experimentally, the composition of the mixture in each amber vial consisted of respective volumes of a 1 mL cell culture in aqueous media, 200 µL n-alkane, and 100 µL BEHP. For the 2LPS, organic solvent and substrate (phase I) and aqueous media (phase II) are immiscible. The system was at thermodynamic equilibrium when the Gibbs free energy state was minimised at constant pressure and temperature, limiting the interactions of 'unlike' molecules and separating the water-rich and hydrocarbon-rich phases. The process flowsheet of the 2LPS was simulated at 20˚C and one atmospheric pressure. UNIF-LL predictions were performed at 25˚C to ensure that temperature does not significantly change predicted activity coefficients for a fair comparison to previous literature with a specified temperature of 25˚C.

Process flowsheets comprised component streams and unit operation devices (Figure 3.2). The aqueous phase was defined as the liquid phase with a higher mole fraction of water (water-rich phase). Two streams (ALKANE and BEHP) were connected to a mixer (MIXER1), allowing oxygenation of the organic phase stream. One stream consisted of n-alkane (ALKANE) and flowed into the mixer at a flow rate of 0.0002 $L{\cdot}min{-}1$ and the other consisted of BEHP organic solvent carrier (BEHP), and flowrate was 0.0001 $L{\cdot}min{-}1$. These flow rates were selected to mimic the aqueous-to-organic phase ratio of 1:0.3 and the substrate-to-organic solvent ratio of 2:1.

The ALKANE and BEHP streams combined in MIXER1 to form the organic liquid phase (phase I), an outlet stream named ORG1. ORG1 outlet stream and a stream consisting of only distilled water (AQ) (phase II) at a flow rate of 0.001 $L{\cdot}min{-}1$ was connected to a mixer (MIXER2) to form two-liquid phases. The two-liquid phases in MIXER2 were passed through an outlet stream (LLE) and sent into a single-stage separator (DECANTER) to separate the two liquid phases based on the UNIF-LL model. The DECANTER split the two-phase liquids by minimising the Gibbs free energy of the system. The resultant streams LIQ1 and LIQ2 represented the 2LPS at equilibrium. The liquid phase that resulted in the higher mole fraction of pure water was denoted as the aqueous phase, and the liquid phase that consisted of primarily organic co-solvent (BEHP) and alkane was denoted as the organic phase. The same process flowsheet was used to determine the impact of different biocompatible solvents (dimethyl sulfoxide, ethyl oleate, hexadecene) on the substrate availability in the aqueous phase as well without cosolvent.

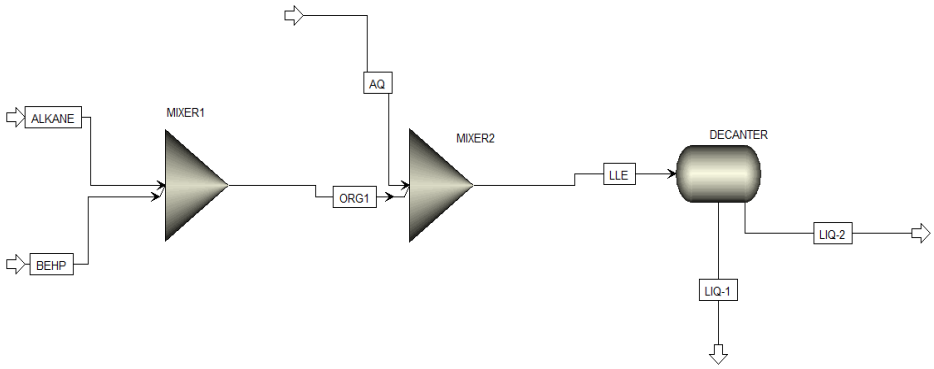


Figure 3.2. Aspen Plus UNIF-LL process flowsheet model of the 2LPS simulating experimental laboratory methods without whole cell biocatalyst. Mole fractions and thermodynamic activity coefficients of alkane in two immiscible phases at minimal Gibbs free energy were provided as an output (LIQ1 and LIQ2). The input stream labelled "ORG1" constitutes mole fractions of the BEHP carrier solvent and alkane (organic phase) previously mixed in MIXER1. The input stream labelled "AQ" constitutes dH2O (aqueous phase). Mixing these two streams (ORG1 and AQ) formed a 2LPS in MIXER2. The outlet stream labelled "LLE" forms the resultant two-liquid phase stream.

3.5.1.1 Validation of UNIF-LL determined activity coefficients using theoretical calculations based on experimentally determined alkane solubility

The UNIF-LL process simulation was validated by theoretical calculations using previously determined experimental solubility of substrates (Abraham et al., 2017). For a pure component at standard state, γ is equal to one and depends on the mole fraction x. At low concentrations (c) of a component in water, the mole fraction x can be derived from c (Equation 3.14) (Ferrario et al., 2018). The mole-based concentration of pure liquid water is 55.5 M at standard state.

$$x = c/55.5\ M \qquad \textit{Equation 3.14}$$

In systems with compounds of poor water solubility, it was assumed that $x \ll 1$, $and\ \gamma$ was independent of x. The solubility limit of a component (c^{sol}) could be used to derive activity coefficients at infinite dilution (Equation 3.15). Activity coefficients obtained from Equation 3.15 were used to validate activity coefficients predicted by UNIF-LL (Banerjee, 1985, Ferrario et al., 2018).

$$c^{sol} = 55.5\ M/\gamma^{\infty} \qquad \textit{Equation 3.15}$$

Parity plots were constructed to determine the extent of correlation between the theoretically calculated or predicted $\gamma\infty$ values and experimentally determined values by Abraham et al. (2017).

3.5.2 Determination of thermodynamic activity-based MM constant (K_m^a)

K_m^a constants of enzyme-substrate (ES) were calculated using Equation 3.16 (adopted by Ferrario et al. (2018) by using previously determined activity coefficients by UNIF-LL and theoretical calculation (Equation 3.15), respectively, and NLR determined K_m^c constants.

$$K_m^a = \gamma^{\infty} \cdot K_m^c/55.5\ M \qquad \textit{Equation 3.16}$$

3.6 The impact of substrate availability on reaction kinetics

Substrate availability in the aqueous phase was determined, which allowed a greater understanding of the impact of substrate concentration on the reaction rate and kinetics of whole cell biohydroxylation on *n*-octane, *n*-nonane, and *n*-decane, respectively. Substrate partitioning behaviour was predicted by attaining mole fractions in each liquid phase at thermodynamic equilibrium using the UNIF-LL process simulations. Partitioning of each alkane in the 2LPS were calculated to evaluate the impact of substrate solubility on substrate availability on reaction kinetics. Aqueous phase substrate availability was related to substrate solubility, alkane chain length, and whole cell biocatalytic performance. UNIF-LL process simulations were used to predict substrate behaviour in 2-liquid-phase systems (2LPSs) by changing the organic co-solvent as well to observe the potential impact of different cosolvents on substrate concentration in the aqueous phase.

3.6.1 Partition coefficients, distribution coefficients, and phase equilibrium constants

The respective mole fractions (x) of alkane in the organic (α) and aqueous phase (β) predicted by UNIF-LL-based Aspen Plus V12 process simulations were used to calculate partition coefficients (K_p) each alkane, respectively, at phase equilibrium (Speight, 2018). The partition coefficient (K_p) is defined as the ratio of mole fractions of the solute component in the two respective phases at thermodynamic equilibrium and measures the escaping tendency of a solute to a particular phase in a liquid-liquid (LL) phase system (Equation 3.17).

$$K_p = \frac{x_\alpha}{x_\beta} \qquad \textit{Equation 3.17}$$

3.6.2 Alkane substrate availability in the aqueous phase at thermodynamic equilibrium

UNIF-LL process simulations were used to calculate the alkane concentration in the aqueous phase. The amount of alkane in the aqueous phase was calculated by multiplying the UNIF-LL predicted mole fraction of alkane in the aqueous phase ($x^{aq}_{alkane\ UNIF-LL}$) by the total amount of alkane placed in the experimental system ($n_{experimental\ total\ alkane}$) (Equation 3.18).

$$n^{aq}_{alkane\ UNIF-LL} = x^{aq}_{alkane\ UNIF-LL} \times n_{experimental\ total\ alkane} \qquad \textit{Equation 3.18}$$

The alkane concentration in the aqueous phase was calculated by assuming that the aqueous phase volume remained constant over time (Equation 3.19).

$$c^{aq}_{alkane\ UNIF-LL} = \frac{n^{aq}_{alkane\ UNIF-LL}}{volume_{\ aq}} \qquad \textit{Equation 3.19}$$

The UNIF-LL predicted aqueous phase concentrations were validated by comparison to previously determined solubility data determined by Riddick et al. (1986).

4 Results and discussion

The kinetics of the whole cell bioconversion of alkanes into primary alcohol in a two-liquid-phase-system were investigated. Two enzymes were selected and expressed in whole cell biocatalysts: CYP153A6 and CYP153A13. Both enzymes have previously demonstrated promising activity for the hydroxylation of medium chain alkanes into primary alcohols. One major difference between the two enzymes is that CYP153A6 is expressed as a multicomponent system with separate electron transfer proteins, whereas CYP153A13 is expressed as a fusion system. To investigate the potential biotechnological challenges, a part of the study aimed to understand the potential benefits or limitations that may arise when using multicomponent cytochrome P450 (CYP) systems compared to fusion CYPs. Chapter 4.1 details the characterisation of the expression and activity of the two CYPs and explores the impact of enzyme expression on biocatalytic performance.

Thereafter, the major focus of the study was performed in which the kinetics of the CYPs between alkanes of different solubility (C_8, C_9 and C_{10}) were tested to determine the potential impact of solubility on the rate of the reaction due to limited access to the enzyme contained within the whole cell in the aqueous phase. The enzyme-substrate affinity could also potentially limit the rate of the reaction and therefore the substrate scope of the whole cell biocatalysts were explored as well. This was done by measuring the rate of the reactions experimentally and by coupling these results to thermodynamic analyses, the true enzyme-substrate affinity could be revealed and effective substrate concentration available to the whole cell biocatalyst residing in the aqueous phase.

4.1 Whole cell biocatalyst preparation and characterisation of expression and activity

The aim was to determine the success of preparing CYP153A13 whole cell biocatalyst and characterise its expression and activity. The turnover number (TON), volumetric alkane bioconversion rate, and biocatalyst efficiency (*viz.* Table 2.2 in Literature Review Chapter 2.4.1) of prepared CYP153A13 whole cells were investigated to evaluate its approximate functionality and performance as a catalyst for the hydroxylation of medium chain alkanes.

4.1.1 Initial characterisation of CYP153A13 expression and activity

4.1.1.1 Heat shock transformation of E. coli BL21 with the CYP153A13 gene

Recombinant organisms carrying the CYP153A13 gene on the pET-28b(+) plasmid, were prepared using heat shock transformation (Appendix 7.1). Successful insertion of the CYP153A13 into the *E. coli* host was confirmed by performing a restriction enzyme diagnostic double digest of the pET-28b (+) vector containing the gene insert, using restriction enzymes HindIII and NdeI. Correctly sized fragments of the digested plasmid were viewed by agarose gel electrophoresis (Figure 4.1). The presence of the CYP153A13 gene was confirmed by the band seen at around 2 kb pair in lanes C, D, and E.

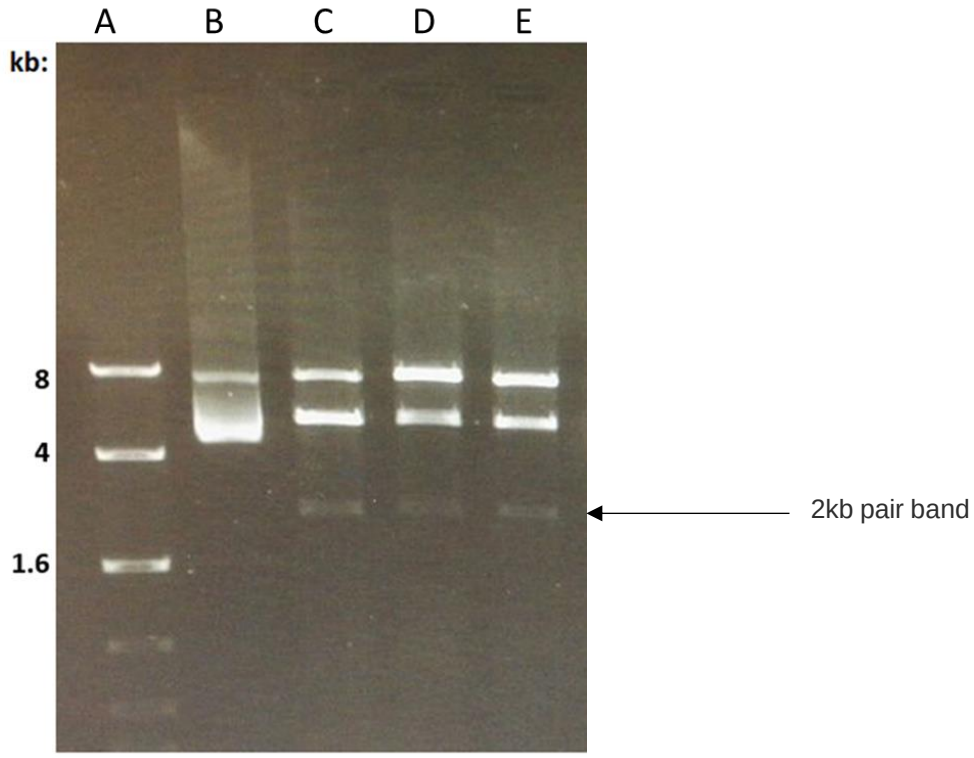


Figure 4.1. Agarose gel electrophoresis of restriction enzyme digested and isolated CYP153A13 DNA plasmids, recovered from the recombinant *E.coli*. The size of DNA fragments was shown by kb, denoting kilobases on the left. Right above the diagram shows the lane labels, which are as follows: (A) Molecular DNA ladder, (B) Undigested plasmid DNA, (C to E) Plasmid DNA digested with HindIII and NdeI.

4.1.1.2 Characterisation of CYP153A13 enzyme expression using SDS-PAGE gels and CO-difference spectral analyses

In this section, the expression, purification and characterisation of CYP153A13 is presented. CYP102A1 mutant is a well-characterised model fusion CYP, hence it was purified alongside CYP153A13 to ensure satisfactory validated purification and characterisation procedures. The CYP102A1 mutant is well known for the selective hydroxylation of the sub-terminal carbon to produce secondary alcohols using NADH, while CYP153A13 shows selectivity for the terminal carbon to produce primary alcohols.

CYP153A13 and CYP102A1 mutant were expressed as recombinant proteins in an autoinduction medium at 20˚C and 200 rpm. After 48 hours of cultivation, the cultures were harvested to recover and purify these enzymes using his-tags. The hexa-histidine-tag was fused to the N-terminal of the proteins to enable purification by using immobilised nickel metal affinity chromatography followed by characterisation using SDS-PAGE gel analyses. SDS-PAGE gel stained with Coomassie® brilliant blue R-250 and heme staining were used to determine the presence and functionality of CYP153A13 and CYP102A1 mutant, as described in Chapter 3.1.4. CO-different spectral analyses were used to detect the concentration of CYP153A13 alongside CYP102A1 mutant, as described in Chapter 3.2.2.

For SDS-PAGE gel analysis, a volume of 10 μL of sample containing CYP153A13 cleared lysate (lane 2), desalted CYP153A13 (lane 3) and dialysed CYP153A13 (lane 4). The same volume of purified CYP102A1 mutant samples were loaded into lane 6 (cleared lysate), lane 7 (desalted) and lane 8

(dialysed) (Figure 4.2A). In lane 5, empty vector crude lysate was added as a control. Purified proteins were further concentred by respectively desalting using a desalting column and dialysed overnight. Both desalted and dialysed samples were concentrated further using ultracentrifugation (Chapter 3.1.5). Figure 4.2B illustrates a heme-stained gel loaded with purified and dialysed CYP153A13 in lane 4 and cleared lysate in lane 2.

Full-length CYP153A13 was displayed at the 95 kDa band in lanes 3 and 4 in the gel shown in Figure 4.2A and lane 4 of the gel in Figure 4.2B, as indicated in the solid circles. In lanes 7 and 8, full length CYP102A1 mutant was observed at the 120 kDa band, similarly observed at 119 kDa in a previous study (Whitehouse et al., 2012). The findings demonstrated in Figure 4.2A indicated that the CYP153A13 and CYP102A1 enzymes were expressed in full length at 95 kDa and 120kDa, respectively. The heme stained gel indicated that the purified dialysed CYP153A13 sample in lane 4 showed proper expression of the enzyme's heme-group.

The cleared lysate CYP153A13 samples in lanes 2 of both gels in Figure 4.2A and B were too dilute to visualise as the amount of P450 expressed per litre bioreaction mixture (BRM) was low (*viz*. CO-difference spectra in Figure 4.3B). This low concentration of CYP153A13 was especially emphasised when comparing the intensity and size of the protein band to that of the cleared lysate of the CYP102A1 mutant in lane 6 of Figure 4.2A. Furthermore, bands in lane 7 and 8 representing the desalted and dialysed CYP102A1 mutant samples had a higher band intensity at the full-length band at 120 kDa compared to the band intensity of CYP153A13 at the full-length band at 95 kDa. This indicated that the concentration obtained from full-length CYP102A1 mutant purification was higher than that of CYP153A13.

Interestingly, the CYP153A13 samples in lanes 3 and 4 showed a contaminant protein which appeared at the 67 kDa band (dotted circles in Figure 4.2) which was in agreement with reports by Bordeaux et al. (2011). Using LC/MS/MS, Bordeaux et al. (2011) identified the protein at 67 kDa as the CYP153A13a heme domain from *Alcanivorax borkumensis* (also known as P450balk), suggesting that the CYP153A13 enzyme underwent partial cytoplasmic proteolysis by trypsin digestion. Bordeaux et al. (2011) suggested that a host cytoplasmic protease cleaves the peptide linker between the CYP and reductase domains. The CYP102A1 mutant also displayed an extra unexpected band around 55 to 67 kDa in lanes 7 and 8, demonstrating a certain extent of proteolytic cleavage. As reported by Fiorentini et al. (2018), the evidence obtained in this study confirmed the susceptibility of CYP153 fusion enzymes to unwanted cytoplasmic proteolytic cleavage.

The evidence in Figure 4.2 suggested that *E. coli* is not an ideal expression host for proteins that are either complex, large, contain disulphide bonds, or require post-translational modifications, substantiating findings by Choi et al. (2006). Fusion proteins often suffer from instability due to proteolytic cleavage by *E. coli* host proteases (LaVallie and McCoy, 1995). Still, *E. coli* is an expression host that is amenable to genetic manipulation and easy to handle in the lab. Should CYP153A13 prove to be a biocatalyst of commercial interest, approaches to avoid proteolytic cleavage should be explored. The evidence obtained in this study in which expression of CYP153A13 is lower than that of CYP102A1 mutant is consistent with the study by Jung et al. (2016), Jacobs et al. (2022) and Bordeaux et al. (2014) who also used SDS-PAGE and CO-difference spectral analyses to illustrate that CYP153A13 had a lower functional and soluble expression compared to other CYPs, such as CYP153A6 and other CYP102A1 mutants (Jung et al., 2016, Jacobs et al., 2022, Bordeaux et al., 2014).

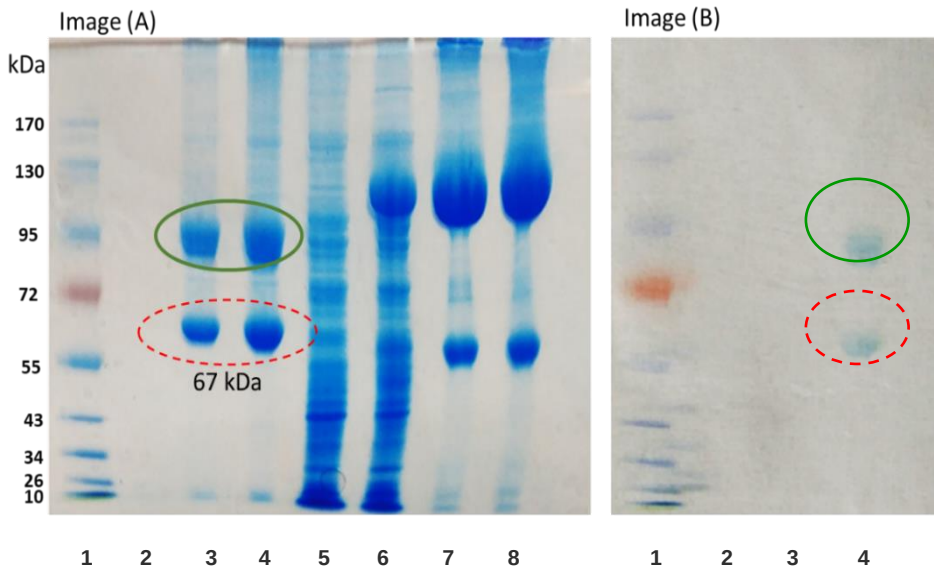


Figure 4.2. Image (A) SDS-PAGE of protein ladder molecular weight marker (lane 1), CYP153A13cleared lysate (lane 2) desalted pure (lane 3) and dialysed pure (lane 4) and of CYP102A1 cleared lysate (lane 6), desalted pure (lane 7) and dialysed pure (lane 8). Lane 5 represents the empty vector crude lysate control. Image (B) Heme stained gel of protein ladder molecular weight marker (lane 1) CYP153A13 cleared lysate (lane 2) and dialysed purified CYP153A13 (Lane 4).

Following visualisation by SDS-PAGE, CO-difference spectra of the cleared lysate, desalted, and dialysed CYP153A13 and CYP102A1 mutant samples were obtained to determine the amount of full-length and functional proteins (Table 4.1 and Figure 4.3). According to CO-difference spectral analyses, CYP153A13 expression and concentration levels were 2.4 to 5-fold lower than CYP102A1 mutant (Table 4.1), confirming the results shown by SDS-PAGE gels above. Crude extracts showed a concentration difference of 3.5-fold between CYP153A13 (3.73 $\mu mol_{P450} \cdot L_{BRM}^{-1}$) and CYP102A1 mutant (12.9 $\mu mol_{P450} \cdot L_{BRM}^{-1}$) (Table 4.1). Similarly, CO-difference data reported by Pennec et al. (2015) showed a four-fold difference between CYP153A13 and the CYP102A1 mutant with concentrations of 1.00 $\mu mol_{P450} \cdot L_{BRM}^{-1}$ and 4.00 $\mu mol_{P450} \cdot L_{BRM}^{-1}$, respectively. Pennec et al. (2015) also used an autoinduction medium for 48 hours at 20˚C and a One-Shot Cell Disruptor and separation of the soluble protein fraction to obtain cell free-extracts. In contrast, the cell lysate sample in the present study underwent an extra purification step using a 0.45 µM filter (Chapter 3.1.5), explaining the four-fold higher CYP153A13 concentration and three-fold higher CYP102A1 concentration than those obtained by Pennec et al. (2015). On comparing the pure dialysed CYP153A13 preparations, which underwent an extra purification step beyond the nickel affinity IMAC chromatography, its concentration is 15.2 $\mu mol_{P450} \cdot L_{BRM}^{-1}$ i.e., 15-fold that of Pennec et al., while the equivalent fraction for CYP102A2 is 9-fold higher.

Although the evidence showed that CYP102A1 mutant was expressed in higher concentrations per litre of BRM, the expression and purification of CYP153A13 was successful as confirmed by SDS-PAGE gel and CO-difference spectral analyses. Since CYP153A13 is a fusion enzyme that has demonstrated selective hydroxylation at the terminal carbon to produce primary alcohols, as opposed to CYP102A1 mutant which does so at the sub-terminal carbon, characterising CYP153A13's expression was crucial because it is a soluble fusion enzyme that can catalysed the primary hydroxylation of linear alkanes, fulfilling a niche in the field of biocatalysis.

Table 4.1. The concentration of CYP153A13 and CYP102A1 mutant in cleared lysate and purified samples measured using CO-difference spectral analyses after 48 hours in autoinduction media at 20˚C

CYP concentration ($\mu mol_{P450} \cdot L_{BRM}^{-1}$)	Cleared lysate	Pure desalted	Pure dialysed
CYP153A13	3.73	5.64	15.2
CYP102A1 mutant	12.9	28.4	36.1

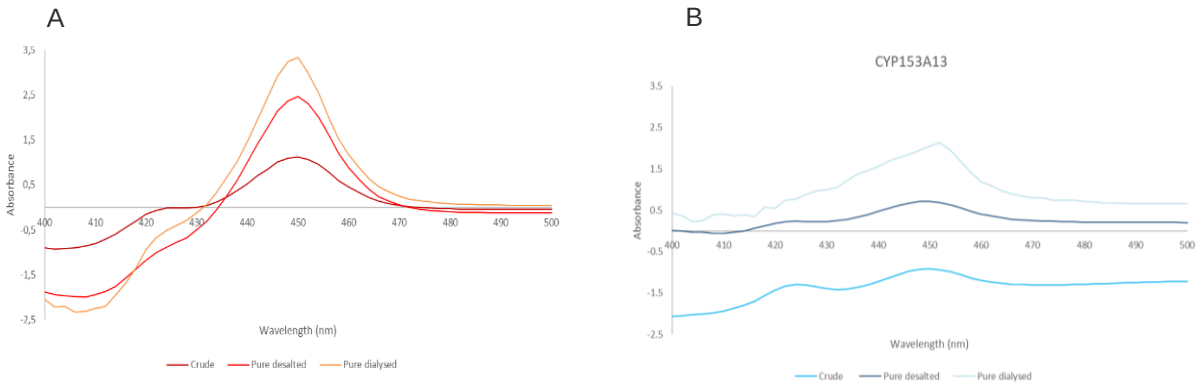


Figure 4.3. CO-difference spectra of crude and purified fusion enzymes (A) CYP102A1 mutant and (B) CYP153A13

SDS-PAGE gels in Figure 4.2 showed that both enzymes underwent proteolysis, but CO-difference spectra results (Table 4.1 and Figure 4.3) confirmed that the full length CYP102A1 showed higher intensity bands (lanes 7 and 8) at 130 kDa compared to the intensity of the full length CYP153A13 band at 95 kDa in lanes 3 and 4. The higher intensity of the bands meant that although CYP102A1 mutant underwent truncation, there was still substantially more of the full-length CYP102A1 mutant compared to CYP153A13 illustrated by the CO-difference spectra results (Table 4.1 and Figure 4.3). These results agreed with Bordeaux et al. (2014), who stated that the CYP102A1 mutant was predicted to have a 23% higher chance of solubility than CYP153A13 when expressed in *E. coli*. The results obtained support the findings by Bordeaux et al. (2014), demonstrating that the expression of soluble and functional CYP153A13 is lower than CYP102A1 (Table 4.1).

To test the activity of the enzyme through biocatalysis, the hydroxylation of alkane was studied according to the method laid out in Chapter 3.1.5. Reactions were carried out at 20˚C at 200 rpm in 60 mL amber vials containing 300 μL of either cleared lysate, desalted or dialysed samples, along with 100 mM glucose, 5 U glucose dehydrogenase· mL^{-1} and 200 mM Tris-HCl at ph 7.4. After 24 hours of the reaction, the primary alcohol produced was analysed using GC after ethyl acetate extraction as described in the method in Chapter 3.2.2. CYP153A13 extracts showed no in vitro activity on *n*-octane due to a lack of expression of the full-length functional enzyme (Appendix 7.7). Crude, desalted, and dialysed CYP102A1 mutant extracts, however, showed turnover numbers (TONs) of 57.7, 12.8, and 9.2 $\mu mol_{2-octanol} \cdot \mu mol_{P450}^{-1}$, respectively (Appendix 7.8). Previously, Pennec et al. (2015) showed that CYP153A13 crude extracts obtained 75% and 50% lower turnover frequencies (TOFs) than CYP153A6 and CYP102A1 mutant after eight hours, respectively. Bordeaux et al. (2014) observed some unproductive CYP153A13 cultures with large variations between batches, which most likely resulted in unsuccessful biohydroxylation of *n*-octane, as observed in the present study.

It is well-recognised that it is challenging to obtain CYP153A13 concentrations sufficient for biohydroxylation of medium chain alkanes (Jung et al., 2016, Bordeaux et al., 2014, Pennec et al., 2015). Although fusion enzymes have shown superior coupling efficiencies to multicomponent enzymes, this advantage is proposed to be counteracted by a lack of properly folded, full length and functional CYP153A13 expression (Ciaramella et al., 2017, Bordeaux et al., 2011, Jung et al., 2016). However, the specificity of CYP153 and the potential for use of CYP153A13 as a 'single pot' enzyme continues to drive interest in attaining an effective CYP153A13 biocatalytic system through developing understanding of the system constraints to inform approaches to over them. Therefore, it was proposed that use of higher cell density BRMs would be beneficial to obtain high concentrations of CYP153A13 at the desired purity to promote the successful biohydroxylation of *n*-octane by CYP153A13, as presented in Chapter 4.1.2.

4.1.1.3 Determination of CYP concentration in whole cell biocatalysts using CO-difference spectral analyses

In addition to comparing the expression of CYP153A13 to a well-studied CYP, it was desirable to compare the biocatalytic performance of the multicomponent enzyme system to fusion CYPs using medium chain alkanes. The multicomponent CYP153A6 system is well recognised for effective biocatalytic performance on *n*-octane, hence it was selected as the multicomponent enzyme in this investigation. CO-difference spectral analyses and in vivo studies were performed. CO-difference spectral analyses confirmed that CYP153A6 showed superior expression to CYP153A13 in *E. coli* BL21 (Table 4.2). Further, CYP153A6 cultures exhibited a redder appearance after protein induction due to the presence of functional and properly folded heme protein, whereas CYP153A13 did not show a red appearance (Figure 4.4). It should be noted that the ratio of P450 to the reductase component in truncated fusion CYP153A13 may have had an impact on electron transfer and consequently the biocatalytic performance towards medium chain alkanes. The CO-difference spectra method, however, measures the presence of a functional electron transfer reductase protein. Therefore, the CO-difference spectra analyses measures only active CYP153A13 with a functional reductase component.

Table 4.2. CO-difference spectral results of intracellular CYP concentrations obtained in this study compared to results in previous literature.

	CYP concentration ($\mu mol\ P450 \cdot g_{DCW}{}^{-1}$)	
	The present study*	**Literature**
CYP153A13	0.0420 ± 0.0065	Pennec et al. (2015) 0.8▪
		Bordeaux et al. (2014) 0.0339
CYP153A6	0.0859 ± 0.0089	Gudiminchi et al. (2012) 0.5-1.0
		Olaofe et al. (2013) 0.2
		Pennec et al. (2014) 1.2▪

In this study and Olaofe et al. (2013), CO-difference spectra were taken after a 24-hour induction using 0.5 mM IPTG at 20˚C. In Pennec et al. (2015) and Gudiminchi et al. (2012), CO-difference spectra were taken after 48 hours using an autoinduction medium. Bordeaux et al. (2014) obtained CO-difference spectra results after 24 hours of 5 mM IPTG induction at 25˚C.

*Average CYP concentrations calculated from three biological repeats ± standard error of the mean at the 95% confidence interval (n=3) (Appendix 7.3)

▪ Estimated from µmol P450·L $^{-1}_{BRM}$ and cell dry weight estimates provided in the literature

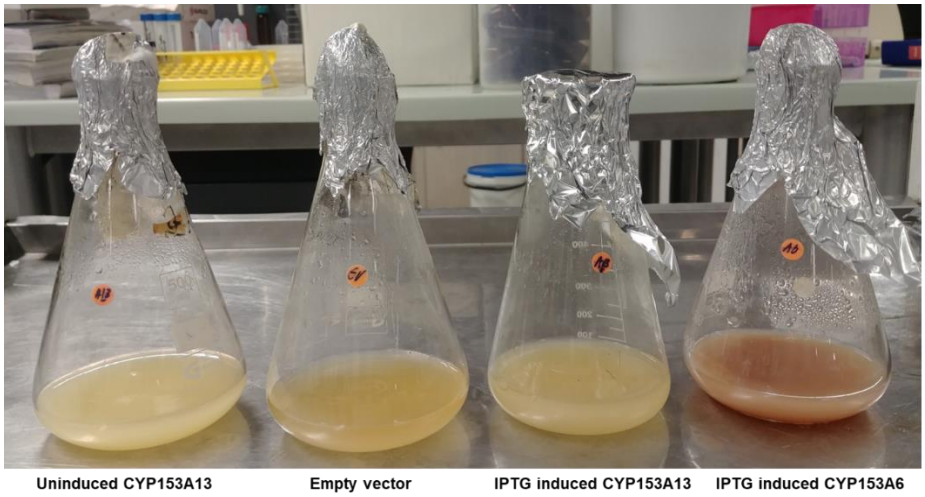


Figure 4.4. Recombinant CYP153A13 and CYP153A6 cells were grown under antibiotic selection with enzyme expression induced by adding 0.5 mM IPTG in mid-exponential phase of growth followed by incubation for 24 hours at 20˚C. CYP153A13 did not show the red colour produced by properly folded and expressed CYP, while the last culture CYP153A6 showed a clear red appearance.

The concentration of CYP153A6 expressed in *E. coli* BL21 was approximately double that of CYP153A13 after induction with 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG) in the mid-exponential phase of growth (OD_{600} of 0.8 to 1), followed by incubation at 20˚C for 24 hours. The intracellular CYP153A6 concentration was 0.0859 ± 0.0089 $\mu mol_{P450} \cdot g_{DCW}^{-1}$, compared to an average of 0.0420 ± 0.0065 $\mu mol_{P450} \cdot g_{DCW}^{-1}$ achieved for CYP153A13. Similarly, Pennec et al. (2015) reported a 1.5-fold higher concentration for CYP153A6 compared to CYP153A13 after 48 hours of expression using an auto-induction medium.

Comparable CYP153A13 concentrations were reported by Bordeaux et al. (2014) following 5 mM IPTG induction for 24 hours at 25˚C, resulting in a concentration of 0.0339 $\mu mol_{P450} \cdot g_{DCW}^{-1}$. Besides a higher IPTG induction concentration used by Bordeaux et al. (2014), similar cultivation and induction conditions were used in the present study. Intracellular CYP153A6 concentrations obtained in this study were more than ten-fold lower than concentrations obtained by Gudiminchi et al. (2012) and Pennec et al. (2015) (Table 4.2). These authors used an autoinduction medium for 48 hours, whereas this study used IPTG induction starting from the mid-exponential growth phase for 24 hours. It was postulated that autoinduction medium may result in higher CYP153A6 concentrations and expression. Previous studies however showed that although autoinduction medium may result in higher CYP concentrations, IPTG induced cultures showed higher specific 1-octanol production rates (Gudiminchi et al., 2012, Olaofe et al., 2013).

4.1.2 Assessment of in vivo biohydroxylation activity using High Cell Density Cultures of whole cell biocatalysts

The high cell density culture (HCDC) approach inspired by Bordeaux et al. (2014) was used to investigate the potential gain in biohydroxylation activity by increasing the cell density of the BRM and thereby increasing the total mole amount of full-length functional CYP153A13. This would increase the active sites available for the primary biohydroxylation of linear alkanes.

Recombinant CYP153A13 and CYP153A6 cells were grown under antibiotic selection with enzyme expression induced by adding 0.5 mM IPTG in the mid-exponential phase (OD_{600} of 0.8 to 1) of growth. After 24 hours of induction, HCDCs and low cell density cultures (LCDCs) were prepared by resuspending the cell pellet in phosphate buffer at pH 7.2 to obtain desired BRM cell densities (Table 4.3). The amount of active CYP in each equivalent volume of BRM was measured using CO-difference spectra before the reaction started, after which 200 μL of *n*-octane and 100 μL BEHP carrier solvent was added to each amber vial, respectively, to allow the biohydroxylation reaction to proceed. After 24 and 48 hours of reaction, 1-octanol concentrations in each BRM were determined using GC (Chapter 3.1.6.2 to 3.2.3). To verify that all product formation was due to enzymatic catalysis, a separate experiment was performed to illustrate that negative control BRMs (uninduced and empty vector cultures) did not show alcohol production (Refer to Section 7.6 in Appendix).

4.1.2.1 Comparison of high and low cell density whole cell biocatalyst cultures

High cell density cultures (HCDCs) and low cell density cultures (LCDCs) were prepared to obtain cellular concentrations of 77.61, 3.449, 86.71 and 3.852 $g_{DCW} \cdot L^{-1}$ for CYP153A6 HCDC, CYP153A6 LCDC, CYP153A13 HCDC and CYP153A13 LCDC, respectively (Table 4.3). Cultures were induced at mid-exponential phase followed by incubation for 24 hours at 20˚C. Expectedly, HCDCs of CYP153A6 and CYP153A13 BRMs possessed a higher total active CYP concentration per litre of BRM (Table 4.3). As expected, the intracellular CYP concentration (μmol_{P450} present per g_{dcw}) was the same for the HCDCs and LCDCs (Table 4.3) with values of 0.193 $\mu mol_{P450} \cdot g_{DCW}^{-1}$ and 0.061 $\mu mol_{P450} \cdot g_{DCW}^{-1}$ for CYP153A6 and CYP153A13, respectively. The CYP153A6 HCDCs and LCDCs showed a three-fold

larger intracellular CYP concentration than the CYP153A13 HCDCs and LCDCs. Results in Table 4.3 demonstrated that HCDCs did indeed allow an increase of active, properly folded, and full-length P450 concentrations, supporting the findings by Bordeaux et al. (2014), Olaofe et al. (2013) and Gudiminchi et al. (2012). These results showed that HCDCs would result in a higher volumetric productivity of primary alcohol because a higher mass of catalyst was present in an equivalent volume of BRM compared to LCDCs.

Table 4.3. Concentrations of CYP153A6 and CYP153A13 in HCDC and LCDC whole cells following a 24 h induction at 0.5 mM IPTG and 20˚C

	CYP153A6		CYP153A13	
	HCDC	LCDC	HCDC	LCDC
Cell density ($g_{DCW} \cdot L^{-1}$)	77.61	3.449	86.71	3.854
CYP concentration ± SE 95% CI ($\mu mol_{P450} \cdot {L_{BRM}}^{-1}$)	14.95 ± 0.062	0.664 ± 0.062	5.310 ± 0.038	0.236 ± 0.038
Intracellular CYP concentration ($\mu mol_{P450} \cdot g_{DCW}^{-1}$)	0.193	0.193	0.061	0.061

The turnover number (TON), defined as the amount of 1-alkanol produced per unit of active CYP enzyme present in the system, was determined after 24- and 48-hour of reaction post induction. This shed insight into the ratio of the amount of 1-alkanol obtained via bioconversion to the number of CYP active sites available to partake in the bioconversion. Therefore, TON depends on CYP concentration, reaction mechanism, and enzyme-substrate (ES) affinity where reactants are in excess.

After 48 hours, the TONs of CYP153A6 and CYP153A13 LCDCs were 2-fold and 7-fold higher than their HCDC counterparts. In line with the 22-fold increase in biomass and consistent concentration of CYP per gram biomass, the amount of active CYP in HCDCs per litre of BRM was much higher than in LCDCs. Therefore, more active sites were available for *n*-octane bioconversion per unit volume in HCDCs. The superior TON for the LCDCs relative to the HCDCs may result owing to the increased cell density leading to a higher requirement of oxygen, glucose, and glycerol, and this not being fully met. The higher volumetric productivity of the HCDCs accompanying this is clearly demonstrated in Figure 4.5 B)

The TONs of CYP153A6 and CYP153A13 LCDCs were similar in value after 24 hours, which implied that both enzymes have similar abilities to biohydroxylate alkanes but attributed to the unique characteristics of each enzyme (Figure 4.5A). For example, the CO-difference spectral analyses in Chapter 4.1.1.3 demonstrated that active CYP153A6 was expressed in higher concentrations than CYP153A13, resulting in higher active site availability for bioconversion. On the other hand, CYP153A13 is a fusion protein improving the electron efficiency of the biocatalytic reaction compared to the CYP153A6 multicomponent system. By 48 hours, the CYP153A6 LCDC had a 3-fold higher TON than the CYP153A13 LCDC, which illustrated the lower levels of CYP153A13 expression in *E. coli* outweighed the benefits of superior electron coupling efficiencies. The changes in TON for CYP153A6

and CYP153A13 from 24 to 48 hours are shown in Figure 4.5A. Both enzymes reach the same TON at 24 hours, at 48 hours, CYP153A6 shows a 3-fold higher TON than CYP153A13. The frequency of contact between the substrate and enzyme affects how quickly the enzyme converts substrate into product. The substrate is consumed and becomes less available as the reaction proceeds. The frequency of contact between the enzyme and substrate is hindered by this decline in substrate availability. Since CYP153A13 shows lower levels of expression than CYP153A6, it is likely that by 48 hours, when the substrate has been consumed and less readily available, CYP153A13 does not come into contact with the substrate as frequently. As a result, CYP153A13 showed a lower TON than CYP153A6 at 48 hours. On the contrary, the TON of CYP153A6 increases from 24 to 48 hours because there is sufficient concentration of the enzyme to facilitate the biocatalytic reaction. It could also be that CYP153A6 has a higher affinity towards *n*-octane than CYP153A13.

The lack of frequency of contact between CYP153A13 and substrate compared to the frequency of contact between CYP153A6, and substrate was exhibited by the 9-fold TON achieved by the CYP153A6 HCDC compared to CYP153A13 HCDC at 48 hours. As stated previously, the lack of frequency of contact between CYP153A13 and substrate was due to its low expression levels.

These results suggested that if CYP1531A13 was adequately expressed, it could have a potential for the hydroxylation of *n*-octane. Although the CYP153A13 system showed a 2.8-fold lower expression of the full-length functional enzyme that the CYP153A6 system, the TON values obtained for the CYP153A13 LCDC were comparable to the CYP153A6 LCDC, indicating compatibility between CYP153A13 and *n*-octane. This supports evidence presented by Bordeaux et al. (2014), who found that although CYP153A13 shows low expression levels, it is still able to catalyse the primary hydroxylation of *n*-octane.

The maximum achievable volumetric production rate of octanol by the CYP153A6 LCDC was 20.09 $\mu mol_{1-octanol} \cdot L^{-1} \cdot min^{-1}$ at 48 hours. This was double the volumetric production rate achieved by the CYP153A6 LCDC at 24 hours at 10.97 $\mu mol_{1-octanol} \cdot L^{-1} \cdot min^{-1}$. In contrast, the CYP153A13 LCDCs achieved a maximum volumetric rate of 3.85 $\mu mol_{1-octanol} \cdot L^{-1} \cdot min^{-1}$ at 24 hours but decreased by 34% at 48 hours to 2.55 $\mu mol_{1-octanol} \cdot L^{-1} \cdot min^{-1}$ (Figure 4.5B). Similarly, the CYP153A13 HCDC showed a 35% decrease from 24 to 48 hours. The observed decrease in volumetric production from 24 to 48 hours by CYP153A13 is expected to be in response to depletion of oxygen and nutrient supply at later stages of the reaction. Additionally, because CYP153A13 has been shown to be expressed in low concentrations and because substrate is consumed as the reaction progresses, decreasing its availability, the decreased in TON from 24 to 48 hours may the result of less frequent contact between the enzyme and substrate. The low levels of recombinant CYP153A13 enzymes limited the improvements that an increase in cell density incurs on volumetric 1-octanol production rates compared to the CYP153A6 multicomponent system, in which substantial improvements are observed in volumetric production with an increase in cell density (Figure 4.5B).

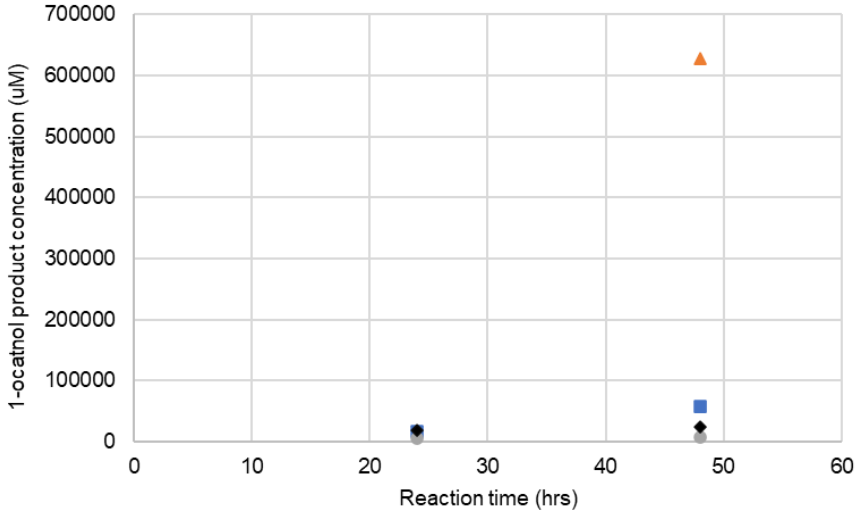
The volumetric rate of the CYP153A6 HCDC at 24 hours was not reported owing to a laboratory error. The CYP153A6 HCDC showed the maximum achievable volumetric rate of 218 $\mu mol_{1-octanol} \cdot L^{-1} \cdot min^{-1}$ at 48 hours and was eleven times higher than its LCDC counterpart while the concentration of active CYP enzyme per volumetric unit of BRM was 22-fold higher (Figure 4.5B). Higher concentrations of active properly folded enzymes present in HCDCs per volumetric unit of BRM similarly impacted the CYP153A13 HCDC, which showed a three-fold higher maximum achievable volumetric rate than its LCDC counterpart.

The results indicated that maximum achievable volumetric rates improved with increased cell density due to increased available full-length and functional CYP concentration. Owing to the consistent specific enzyme concentration, volumetric rates are a function of cell density. CYP153A13 was expressed in lower specific concentrations, which decreased the availability of active site per unit volume compared to CYP153A6 (Figure 4.5). HCDCs were proven to be attractive for superior maximum achievable volumetric primary alcohol production rates, especially if the limitation of oxygen and nutrient supply could be circumvented by implementing a continuous stream of oxygen, glucose, and glycerol to whole cell biocatalysts. Similar findings for HCDCs have been reported by others for biocatalytic systems (Bordeaux et al., 2014, Salehmin et al., 2013, Choi et al., 2006, Glazyrina et al., 2010).

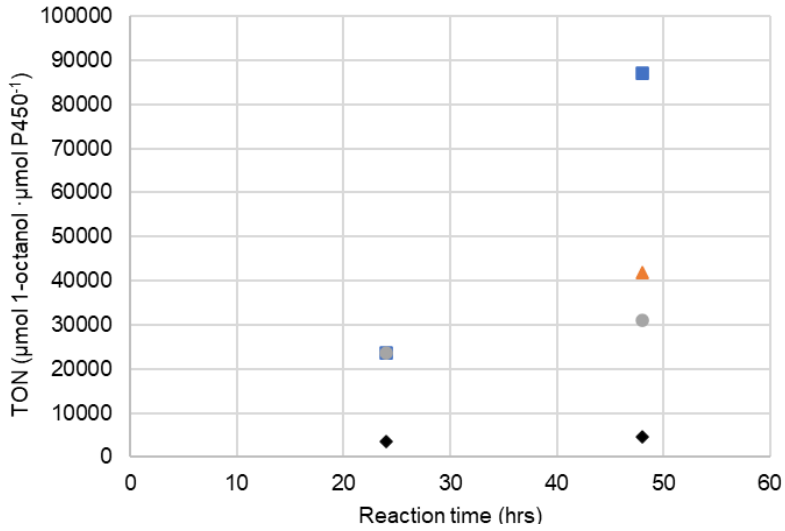
The biocatalytic efficiency was determined at 24 and 48 hours respectively, to provide an indication of the mass of product formed per unit mass of dry cell weight biocatalyst. Although a gram of dry cell weight contains the same amount of active CYP in HCDC and LCDCs (Table 4.3), LCDC cultures have greater oxygen and nutrient availability per cell and hence per molecule of enzyme, increasing the biocatalyst efficiency of each mass unit of cell weight (Figure 4.5C). The CYP153A6 and CYP153A13 LCDCs had higher biocatalyst efficiencies at both 24 and 48 hours than their HCDC counterparts (Figure 4.5C). The biocatalyst efficiency of the CYP153A6 HCDC was 2-fold lower than its LCDC counterpart, with a value of 1054 $mg_{1-octanol} \cdot g_{DCW}^{-1}$ at 48 hours. The CYP153A6 LCDC had the highest biocatalyst efficiency at 2184 $mg_{1-octanol} \cdot g_{DCW}^{-1}$ at 48 hours (Figure 4.5C). Similarly, CYP153A13 LCDC had a biocatalyst efficiency of 248 $mg_{1-octanol} \cdot g_{DCW}^{-1}$ at 48 hours, 7-fold higher than the HCDC at 37 $mg_{1-octanol} \cdot g_{DCW}^{-1}$.

The CYP153A6 LCDC had a 9-fold higher biocatalyst efficiency than the CYP153A13 LCDC at 48 hours. Similarly, the CYP153A6 HCDC showed a 28-fold higher biocatalyst efficiency than the CYP153A13 HCDCs at 48 hours. This resulted from the 3-fold lower availability of CYP153A13 enzyme per gram of dry cell weight compared to cultures expressing CYP153A6. CYP153A6 may also have higher affinity towards *n*-octane, contributing to the higher biocatalyst efficiencies achieved by CYP153A6 low and high cell density cultures compared to CYP153A13.

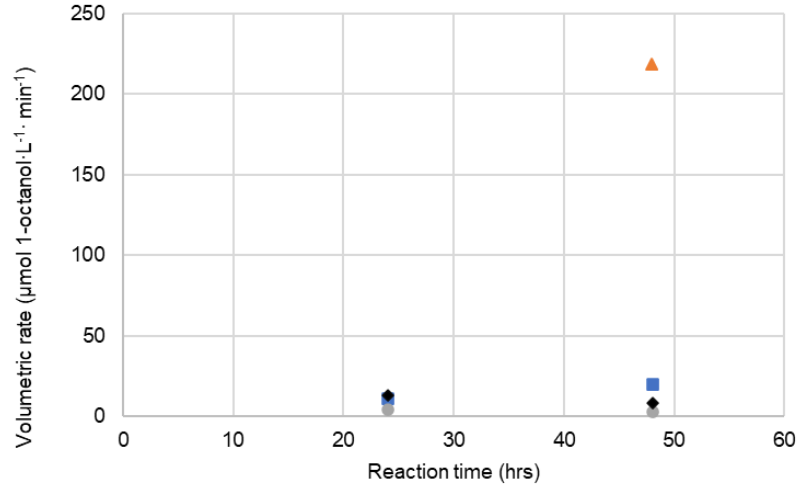
(A)



(B)



(C)



(D)

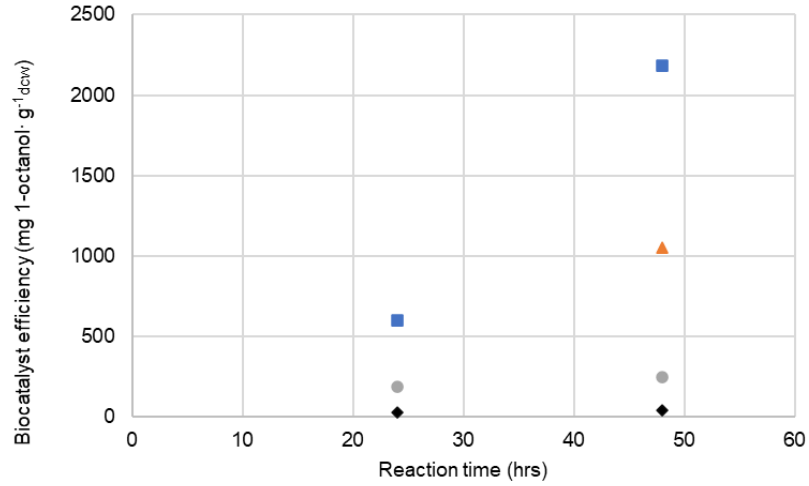


Figure 4.5. (A) Product concentration, (B) TON, (C) Volumetric rate and (D) Biocatalyst efficiency of 1-octanol by HCDC and LCDC CYP153A6 and CYP153A13 recombinant cultures at 20˚C after 24 hours of IPTG induction. The CYP153A6 HCDC data point at 24 hours was not reported because it was not detected.

4.1.3 Concluding remarks

CYP102A1 mutant was used as a comparative model for soluble fusion enzymes to validate experimental approach and thereby CYP153A13 expression was characterised. Overall, CYP153A13 showed a tendency of unwanted, but partial, truncated expression according to SDS-PAGE gel results. CYP153A13 also showed low expression levels which resulted in the absence of *in vitro* biohydroxylation of *n*-octane, however using low cell density cultures, CYP153A13 showed *in vivo* activity on *n*-octane to produce 1-octanol but the activity was lower compared to CYP153A6 cultures. To overcome the low expression of active CYP153A13, high cell density cultures (HCDCs) were used and compared with the CYP153A6 multicomponent system which is known to have sufficient expression levels and the ability to biohydroxylate *n*-octane at the terminal carbon to produce primary alcohols.

The specific intracellular concentrations of CYP153A13 were consistent between low and high cell density cultures, while the specific intracellular concentration for CYP153A6 was 3-fold larger than CYP153A13. Biocatalytic experiments using CYP153A13 at low cell densities showed variable biohydroxylation activity at different times in the reaction. This was dependent on the amount of active CYP153A13 available to contact with substrate, the latter of which also decreases in availability as the reaction proceeds. CYP153A13 HCDCs showed a 3-fold improvement in volumetric production rate because a higher amount of CYP153A13 active site was available for contact with the substrate. Despite the variability, CYP153A13 showed the potential to act as a whole cell biocatalyst for the primary hydroxylation of *n*-octane, which is in agreement with findings by Bordeaux et al. (2014).

Biocatalytic activity was observed for both CYP153A6 and CYP153A13 high and low cell density cultures. The CYP153A6 system showed superior biocatalyst efficiencies and maximum achievable volumetric rates compared to CYP153A13, illustrating the impact of low expression concentration of CYP153A13 compared to CYP153A6. This could have also resulted in superior affinity of CYP153A6 towards *n*-octane compared to CYP153A13. The HCDCs improved the maximum achievable volumetric 1-octanol production rates, but oxygen and nutrient supply appeared limiting compared to the LCDC counterparts. At earlier stages of the reaction at 24 hours, CYP153A6 and CYP153A13 showed the same turnover number. The negative impact of low CYP153A13 expression levels on the biocatalytic performance became evident as substrate was consumed because CYP153A6 showed an increase in turnover number from 24 to 48 hours, while CYP153A13 showed a decrease.

It is likely that CYP153A13 may be a promising catalyst for the primary hydroxylation of *n*-octane. The concentration of active CYP153A13 in the BRM could be increased by optimising its expression strategy using molecular engineering tools. Besides the low expression exhibited by CYP153A13, it was found that substrate concentration available to the whole cell biocatalyst as well as favourable enzyme-substrate affinity may have limited the reaction catalysed by CYP153A13 whole cells. In the light of these findings, the kinetics of CYP153A13 on medium-chain alkanes were further investigated to explore its substrate scope and the potential effects of substrate solubility on biocatalytic 2LPSs.

4.2 Substrate scope of CYP153A6 and CYP153A13 and the impact of substrate solubility on whole cell 2LPS biohydroxylation of C_8, C_9 and C_{10} linear alkanes

CYP153A6 and CYP153A13 are valued for the specificity of their catalysis and their targeting of reaction at the primary carbon; however, to date their recognised substrate range is small. The second experimental study was planned to expand our knowledge of the substrate scope of CYP153A6 and CYP153A13 whole cells and assess the general impact of substrate solubility on whole cell 2LPS biocatalysis. Recombinant CYP153A13 and CYP153A6 whole cells were grown under antibiotic selection with enzyme expression induced by adding 0.5 mM IPTG at the mid-exponential growth phase (OD_{600} of 0.8 to 1) (Chapter 3.1.5.1). After 24 hours of incubation at 200 rpm and 20˚C, the cells were harvested and bioreaction mixtures (BRM) prepared by resuspending the cell pellet in phosphate buffer at pH 7.2 to obtain final cell concentrations of 3 to 5 $g_{DCW} \cdot L^{-1}$. The amount of active CYP was measured using CO-difference spectra before the reaction. Thereafter a volume of alkane (200 µL) and 100 µL BEHP cosolvent, respectively, was added to each amber vial to allow the biohydroxylation reaction to proceed at 20˚C and 200 rpm. Gas chromatography (GC) was used to determine the accumulated product concentration at selected time points from 3 to 72 hours (Chapter 3.2.3).

The turnover frequency (TOF) of the respective enzyme-substrate (ES) combinations for the biohydroxylation of alkanes of different chain lengths and solubility were compared after 24 hours of the reaction post induction (Chapter 4.2.1 and 4.2.2). The chain lengths of alkane (C_8, C_9 and C_{10}) were selected as substrate impacts both the enzyme affinity for the alkane and alkane solubility in the aqueous phase. Alkane solubility decreases with increased chain length (Table 4.4). The biocatalytic activity towards the oxyfunctionalisation of medium-chain alkanes was assessed over 24 and 48 h biocatalytic reaction. Furthermore, to investigate the potential gain of *in vivo* activity on medium-chain alkanes by the fusion enzyme CYP153A13 compared to the CYP153A6 multicomponent enzyme system, the TOF achieved by CYP153A13 was compared to the CYP153A6.

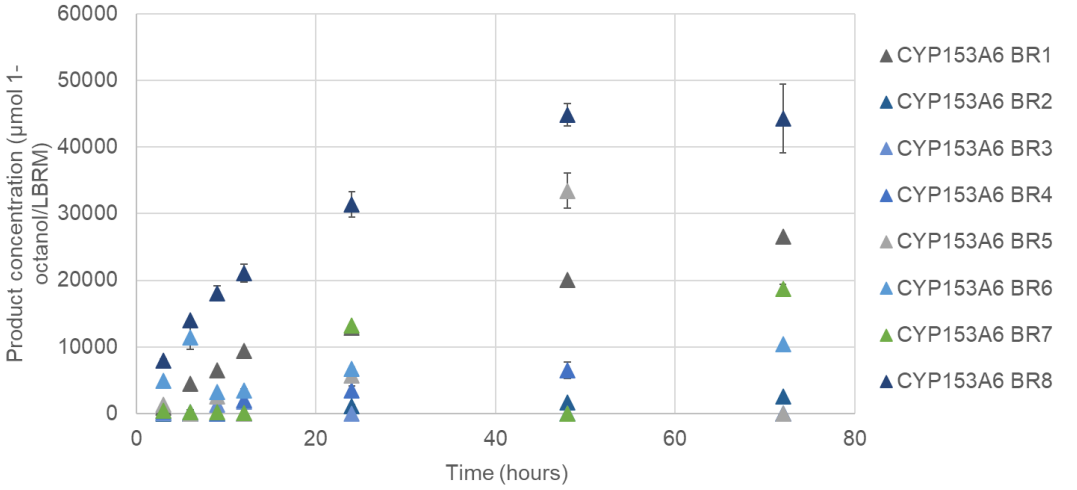
Table 4.4 Solubility of series of alkanes in water at 20˚C, 1 atm (Riddick et al., 1986).

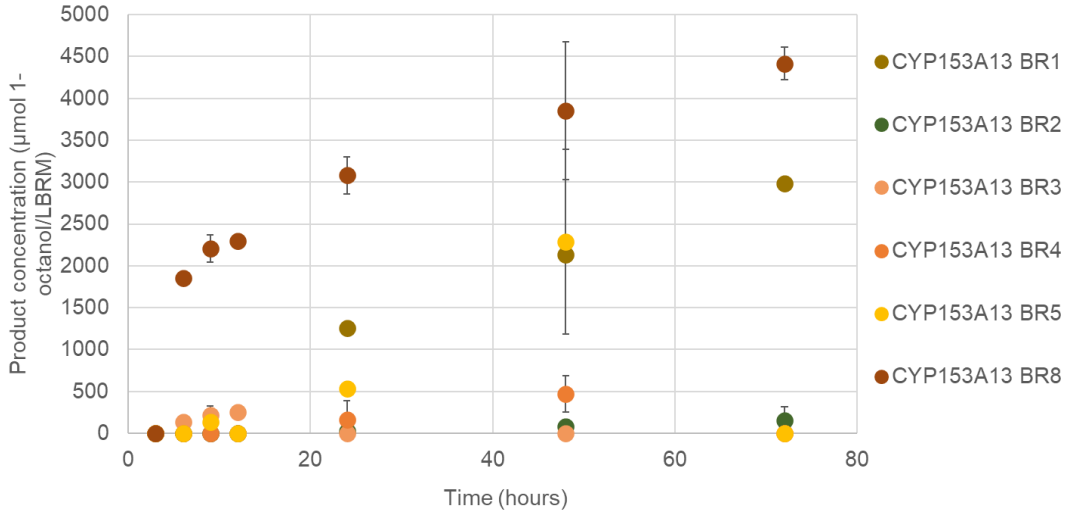
Substrate	*n*-octane	*n*-nonane	*n*-decane
Solubility (μM)	4.00	1.22	0.27

On observing the production of the alcohols with time through biohydroxylation catalysed by CYP153A6 and CYP153A13 in Figure 4.6, substantial variability was observed amongst biological repeats because it can be challenging to precisely regulate variables such as protein expression, cell density, cellular stress, induction conditions, medium composition, aeration, mass transfer, and pH in small-scale cultivation. Bordeaux et al. (2014) also noted the challenge of obtaining reproducible biological repeats in CYP153 biocatalysis. Variability of protein expression occurs even in genetically identical cells (Nordholt et al., 2017). Controlling and precisely regulating the abovementioned variables could be overcome using instrumented large-scale bioreactors that monitor process variables over time using sensors (Bordeaux et al., 2014).

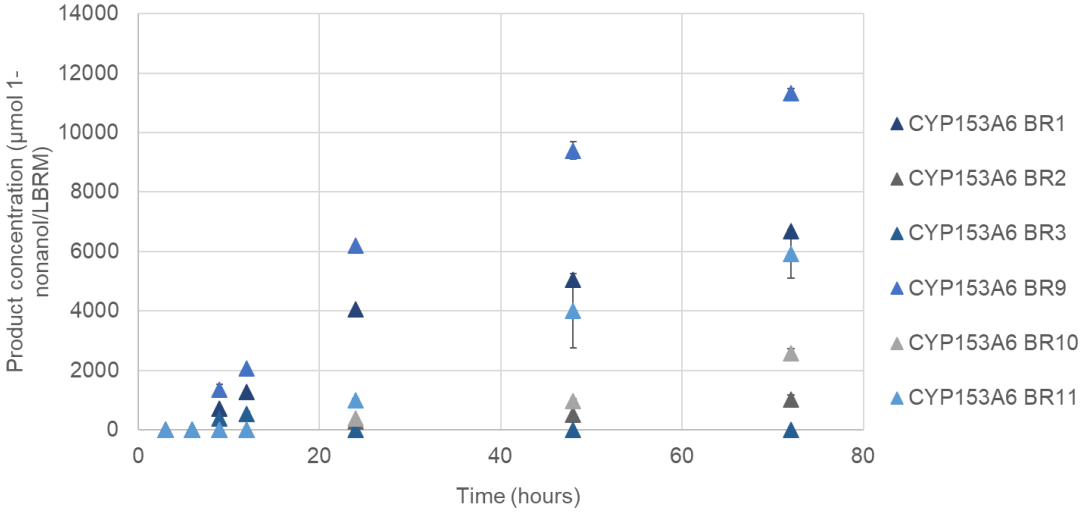
To eliminate the challenge of comparing different alkanes across biological repeats that displayed high variability, analyses of the impact of substrate solubility were compared within a single biological repeat. This analysis is detailed in Chapter 4.2.1.

A


Product concentration (µmol 1-octanol/LBRM)
60000
50000
40000
30000
20000
10000
0
0
20
40
60
80
Time (hours)
CYP153A6 BR1
CYP153A6 BR2
CYP153A6 BR3
CYP153A6 BR4
CYP153A6 BR5
CYP153A6 BR6
CYP153A6 BR7
CYP153A6 BR8


Product concentration (µmol 1-octanol/LBRM)
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
0
20
40
60
80
Time (hours)
CYP153A13 BR1
CYP153A13 BR2
CYP153A13 BR3
CYP153A13 BR4
CYP153A13 BR5
CYP153A13 BR8

B


Product concentration (µmol 1-nonanol/LBRM)
14000
12000
10000
8000
6000
4000
2000
0
0
20
40
60
80
Time (hours)
CYP153A6 BR1
CYP153A6 BR2
CYP153A6 BR3
CYP153A6 BR9
CYP153A6 BR10
CYP153A6 BR11

C

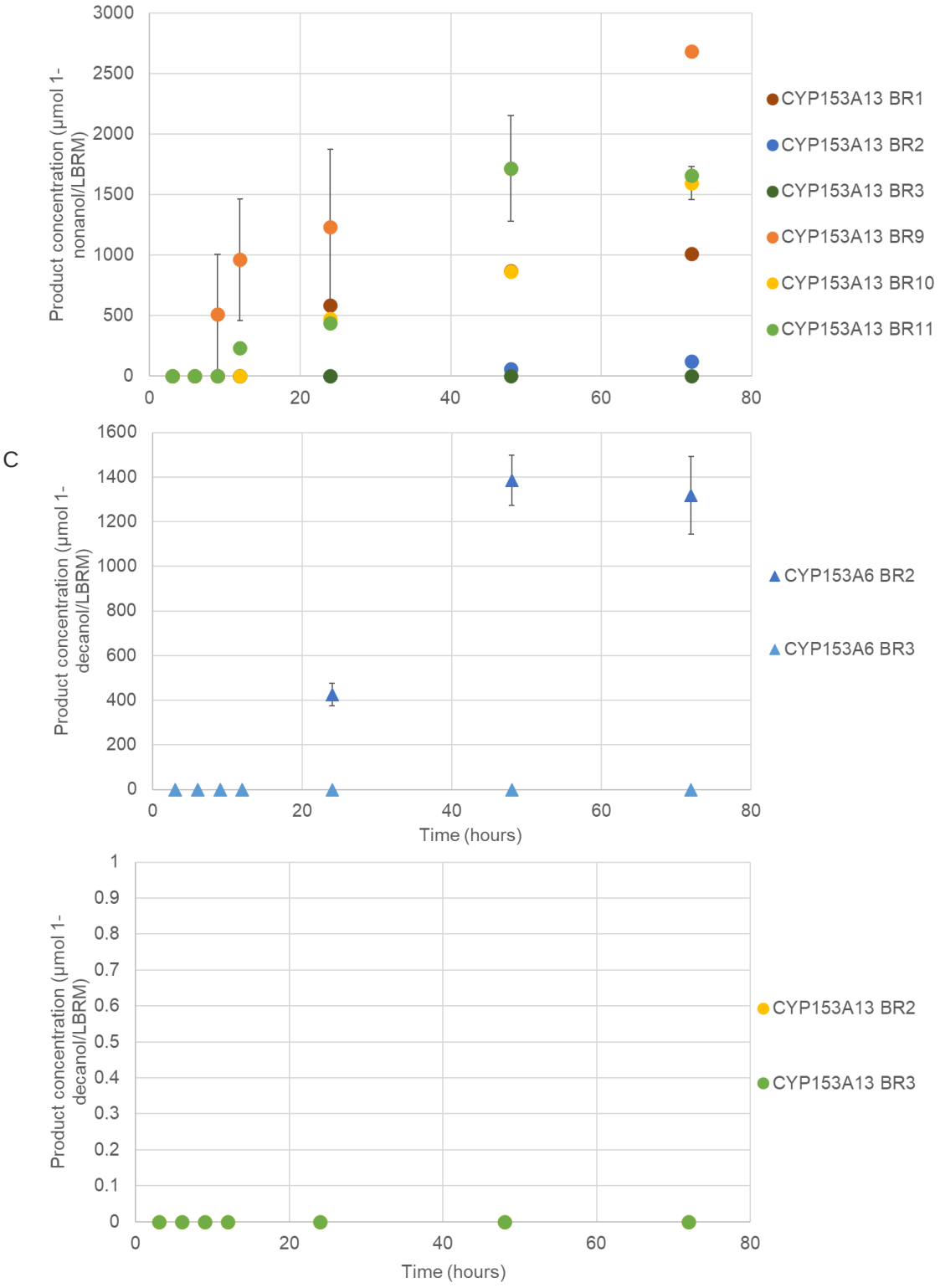


Figure 4.6 Graphs illustrating the variability of the primary product concentration formed over time between biological repeats of (A) 1-octanol, (B) 1-nonanol, and (C) 1-decanol from the biohydroxylation of alkanes in different biological repeats (BR) performed on different days. Error bars represent the standard deviation of the average of a duplicate sample from one recombinant stock culture.

4.2.1 Assessment of the impact of alkane solubility and chain length on TOFs achieved by CYP153A6 and CYP153A13 whole cell biocatalysts

In comparing the impact of chain length and solubility, the data was analysed within each of three biological repeats: BR1, BR2, BR3. BR1 included the measurement of CYP153A6 and CYP153A13 in vivo activity on *n*-octane and *n*-nonane from 3 to 72 hours (Figure 4.7A). CYP153A6 and CYP153A13's *in vivo* activity on *n*-octane, *n*-nonane, and *n*-decane were measured in BR2 and BR3 from 24 to 48 hours and 3 to 12 hours, respectively (Figure 4.7 B & C). It should be noted that there was considerable variability between BR1, BR2 and BR3 and therefore the data was analysed within each biological repeat in isolation, rather than across biological repeats. This could have been due to genetic variation arising from the initial glycerol stocks used for pre-inoculum cultures in different biological repeats BR1, BR2 and BR3. Aliquots of subsamples of microorganisms were regularly frozen over time in multiple experiments, which may have introduced genetic heterogeneity. This may explain the differences observed between the different biological repeats within the study.

In BR1 and BR2, the time it took for shorter-chain alkanes to undergo bioconversion into primary alcohol product was shorter than for the longer-chain alkanes (Figures 4.6 and 4.7). Shorter chain alkanes have a higher aqueous solubility; therefore, shorter chain alkanes have increased availability in the aqueous phase to whole cell biocatalysts and consequently increased the probability of enzyme-substrate (ES) contact (Table 4.4). For example, CYP153A6 whole cells showed the onset of biohydroxylation activity on *n*-octane three hours after reaction initiation, while biohydroxylating *n*-nonane by 9 hours (Figure 4.7A).

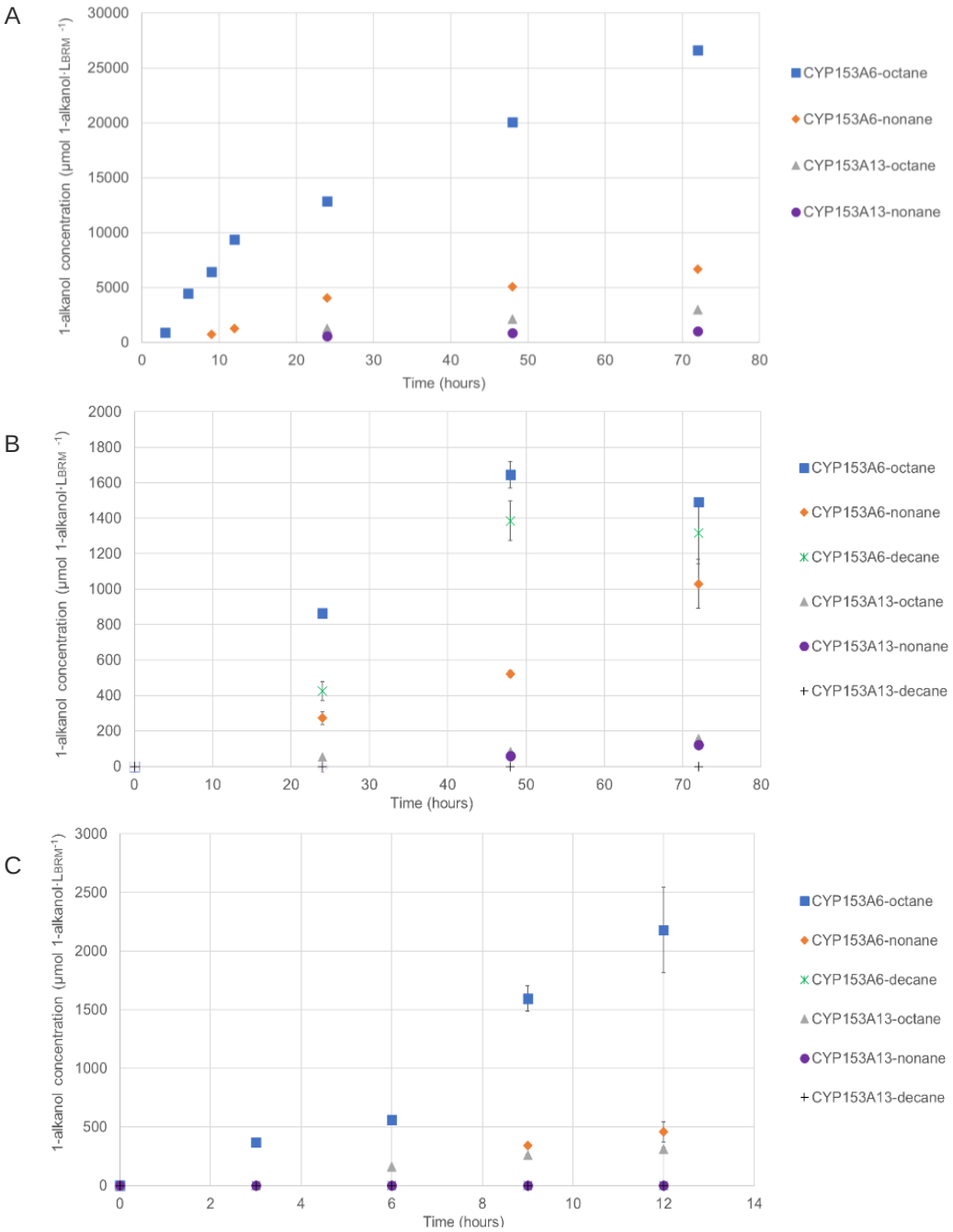


Figure 4.7 Graphs illustrating the 1-alkanol product concentration over time measured using a GC at selected time points to monitor the oxy-functionalisation reaction catalysed (20˚C, 1 atm, 200 rpm) by whole cells containing CYP153A6 and CYP153A13, respectively. (A) BR1 showed the measured 1-alkanol concentration without a duplicate measurement. The concentration of primary alcohol in BR2 (B) and BR3 (C) were measured as an average of duplicates. The error bars represent the standard deviation of the average of a duplicate sample from one recombinant stock culture.

CYP153A13 whole cells showed the onset of activity on *n*-nonane at 48 hours following the reaction initiation in BR2 while showing activity on *n*-octane 24 hours earlier (Figure 4.7B). These findings validated a study by Grant et al. (2011), who attributed the delay of biohydroxylation activity with *n*-dodecane compared to *n*-octane to mass transfer across the cellular membrane or the liquid-liquid (LL) interface. Moreover, a delay was observed in reaction initiation for CYP153A13 cultures, explained by low cytosolic concentrations of CYP153A13 enzyme compared to CYP153A6 and, consequently, the frequency of ES contact. Since the substrate's solubility and the enzyme's expression levels play a role

in the time taken for the reaction to initiate, it was clear that interpretation of biocatalytic data over time is essential to an accurate understanding of enzyme activity on a substrate.

Single-time point analyses can provide insight into enzyme functionality on a substrate and an overview of the impact of substrate solubility on resulting product concentrations (Figure 4.7) (Gardossi et al., 2010). The rate of alkane biohydroxylation reactions of substrates with different solubilities was analysed in terms of their TOF over 24 hours to compare with previous studies (many of which commonly report data using single-time point analyses at 24 hours) and in the cases of delayed reaction initiation, 48 hours (Tables 4.5 to 4.8). The TOF is the derivative of the TON with respect to time and provided insight into the amount of product formed in a unit of time relative to a unit of active enzyme, expressed as $\mu mol_{1-alkanol} \cdot \mu mol_{P450}^{-1} \cdot min^{-1}$.The TOF at 24 hours for each ES pair was determined and compared to previous studies of CYP153 hydroxylation of different alkane chain lengths (Tables 4.5 to 4.8). Following single-time point analyses, the TOFs achieved over time were investigated further in Chapter 4.2.3.

Variability in biocatalytic outcomes was found across several studies, attributed to different reaction conditions, such as the use of whole cells versus purified enzymes and time of measurement. For example, the TOF measured at 8 hours by Pennec et al. (2015) or at 45 minutes by Funhoff et al. (2006), will be, as expected, different from a TOF measured after 24 hours of the reaction. BR1 showed the most similarity to Olaofe et al. (2013) and Guduminchi et al. (2012) who used similar reaction conditions and measured the TOF at 24 hours (Table 4.5). BR2 showed less similarity to BR1 and previous studies by Olaofe et al. (2013) and Guduminchi et al. (2012). As stated previously, Bordeaux et al. (2014) noted the challenge of obtaining reproducible biological repeats in CYP153 biocatalysis. Further, variability of protein expression occurs even in genetically identical cells (Nordholt et al., 2017). Due to the different reaction systems used between the present study and previous literature, the comparisons made are to illustrate general trends of the impact of substrate solubility.

In BR1, the TOF achieved at 24 hours by CYP153A6 whole cells on *n*-octane was 16 $\mu mol_{1-octanol} \cdot \mu mol_{P450}^{-1} \cdot min^{-1}$, while Gudiminchi et al. (2012) and Olaofe et al. (2013) obtained 21 $\mu mol_{1-octanol} \cdot \mu mol_{P450}^{-1} \cdot min^{-1}$ at 24 hours, and Pennec et al. (2015) obtained 9.2 $\mu mol_{1-octanol} \cdot \mu mol_{P450}^{-1} \cdot min^{-1}$ at 8 hours. Funhoff et al. (2006) obtained a much higher value of 60.8 $\mu mol_{1-octanol} \cdot \mu mol_{P450}^{-1} \cdot min^{-1}$ at 45 minutes through using purified enzyme, not whole cell biocatalyst. The TOF achieved by CYP153A6 whole cell hydroxylation on *n*-octane in BR1 was most comparable to reports by Gudiminchi et al. (2012) and Olaofe et al. (2013), who used similar induction and reaction conditions such as IPTG induction for 24 hours in a resting whole cell 2LPS at 20˚C (Table 4.5).

Table 4.5 Alkane biohydroxylation conditions and performance parameters of CYP153A6 whole cells determined in the present study with comparison to literature studies

	The present study (BR1)[1]		Pennec et al. (2015)		Olaofe et al. (2013)	Gudiminchi et al. (2012)	Funhoff et al. (2006)		
Biocatalyst	*E. coli* CYP153A6		*E. coli* CYP153A6		*E. coli* CYP153A6	*E. coli* CYP153A6	Purified CYP153A6		
Growth medium	LB		ZYP-5051 autoinduction		LB	LB	E2 medium		
Induction	IPTG for 24 hrs		Autoinduction for 48 hrs		IPTG for 24 hrs	IPTG for 24 hrs	Not needed- host organism *Pseudomonas Putida* used		
Biotransformation medium	2LPS		2LPS		2LPS	2LPS	2LPS		
Buffer	200 mM sodium phosphate at pH 7.2		200 mM Tris-HCl at pH 8		200 mM sodium phosphate at pH 7.2	200 mM sodium phosphate at pH 7.2	50 mM potassium phosphate at pH 7.4		
Time of measurement (hrs)	24		8		24	24	0.75		
Biotransformation temperature (˚C)	20		20		20	20	30		
Biomass concentration ($g_{DCW} \cdot L_{BRM}^{-1}$)	7.2		5		11	11	n.d		
CYP concentration ($\mu mol \cdot L_{BRM}^{-1}$)	0.55		6.7		0.8	2.0	2.0		
Alkane	C_8	C_9	C_8	C_{10}	C_8	C_8	C_8	C_9	C_{10}
TOF ($\mu mol_{1-alkanol} \cdot \mu mol_{P450}^{-1} \cdot min^{-1}$)	16	5.1	9.2	0.2	21	21	60.8	49.7	13.8

n.d= not determined

1: Please note that BR1 included tests on the activity of CYP153A6 and CYP153A13 whole cells on C_8 and C_9 from 0 to 72 hours. BR2 included tests on the activity of CYP153A6 and CYP153A13 whole cells on C_8, C_9 and C_{10} from 24 to 72 hours. BR3 included tests on the activity of CYP153A6 and CYP153A13 whole cells on C_8, C_9 and C_{10} from 0 to 12 hours.

Table 4.6 Key parameters of CYP153A6 whole cell biohydroxylation of *n*-octane, *n*-nonane, and *n*-decane in BR2. The TOF and biocatalyst efficiency values presented are the average of duplicates (mean ± standard deviation)

Time of measurement (hrs)	24			48		
Biotransformation temperature (˚C)	20			20		
Biomass concentration ($g_{DCW} \cdot L_{BRM}^{-1}$)	8.8			8.8		
CYP concentration ($\mu mol \cdot L_{BRM}^{-1}$)	0.673			0.673		
Substrate	C_8	C_9	C_{10}	C_8	C_9	C_{10}
TOF ($\mu mol_{1-alkanol} \cdot \mu mol_{P450}^{-1} \cdot min^{-1}$)	1.09 ± 0.27	0.28 ± 0.04	0.44 ± 0.05	0.85 ±0.04	0.27 ± 0.01	0.71 ± 0.06

CYP153A6 whole cell biohydroxylation achieved a three-fold higher TOF on *n*-octane than on *n*-nonane at 24 hours (Table 4.5). Similarly, results in BR2 showed a four-fold higher TOF achieved by CYP153A6 whole cells on *n*-octane than *n*-nonane at 24 hours (Table 4.6). As substrate chain length increases and its aqueous solubility decreases, the TOFs achieved by the respective whole cells decrease. Comparing the product concentration and TOF of respective whole cell biocatalysts on different alkane chain lengths demonstrated a higher TOF for shorter chain length (Tables 4.5 and 4.6). This revealed that substrate solubility plays a key role in the effective transport across the cellular membrane and LL interface. Funhoff et al. (2006) achieved a 1.2-fold higher TOF on *n*-octane than *n*-nonane using purified CYP153A6 at 45 minutes, supporting the impact of solubility. The lower impact in their studies with purified enzyme may also suggest increasing limitation of transport of alkane into the whole cell biocatalyst; transport limitation for octane has been clearly shown in the studies of White (2022).

In BR2, after 24 hours, CYP153A6 whole cells achieved a 2.5-fold higher TOF for *n*-octane than *n*-decane, re-illustrating the relationship that a decrease in substrate solubility results in lower achieved TOFs due to lower transport across the cellular membrane and LL interface (Table 4.6). These results were supported by Pennec et al. (2015), who illustrated that permeabilised whole cells expressing CYP153A6 and CYP153A13 showed an increase in activity on *n*-decane compared to non-permeabilised cells, whereas no improvement was observed on *n*-hexane in permeabilised cells compared to non-permeabilised cells. Therefore, the findings in Table 4.5 and 4.6, showing a higher TOF achieved by CYP153A6 whole cells on *n*-octane compared to *n*-decane supports Pennec et al. (2015)'s findings and highlights that a decrease in substrate solubility limits the transport of the substrate across the cellular membrane.

When comparing *n*-nonane to *n*-decane, however, it was illustrated that the relationship between substrate solubility and TOF did not hold true and that the higher solubility of *n*-nonane did not result in a higher TOF (Table 4.6). In fact, the TOF achieved on *n*-nonane was 1.6-fold lower than on *n*-decane

at 24 hours (Table 4.6). This result was unexpected because *n*-nonane is five times more soluble in the aqueous phase than *n*-decane. In contrast, Funhoff et al. (2006) reported that, using purified CYP153A6, they achieved a 3.6-fold higher TOF on *n*-nonane than *n*-decane at 45 minutes. The results therefore show that *n*-decane was more effectively transported across the cellular membrane than *n*-nonane to achieve higher TOFs on *n*-decane (Table 4.6). This may be related to the odd-even effect of liquid alkanes, further discussed in this section.

These findings demonstrated that several factors related to substrate solubility contributed to the TOF achieved by whole cell biocatalysts, potentially including the aqueous alkane concentration, enzyme-substrate affinity, substrate-membrane interactions, substrate-solvent interactions, or the transport of alkanes across the LL interface. It was a potential that CYP153A6 showed higher affinity towards *n*-decane than *n*-nonane, which was investigated further in Chapter 4.3 and 4.4. With regards to transmembrane transport, it was noted that *n*-octane, *n*-nonane, and *n*-decane all have molecular masses of less than 600 Da, which means they are passively transported across the cellular membrane of gram-negative bacteria such as *E. coli* through porins (Wiener and Horanyi, 2011). The results may potentially reflect the odd-even effect of liquid alkanes, described as the alternation of various properties of *n*-alkanes (C_nH_{2n+2}) as a function of carbon content (n) (Dhiman et al., 2022).

The impact of odd- compared to even- chain-length liquid alkanes on transmembrane transport and LL interfacial transport in whole cell biocatalytic 2LPSs should be further investigated. It was conceivably hypothesised that even-chain lengths (*n*-decane) are preferentially transported across either the LL interface of the 2LPS or the cellular membrane of gram-negative bacteria than odd-numbered alkanes (*n*-nonane) (Dhiman et al., 2022). Interestingly, this evidence supports a study by He et al. (2016) showed that the degradation of alkanes by bacteria shows preference for even chain alkanes (He et al., 2016). Alkanes with an even number of carbons show a higher ordered periodic packing efficiency and isotropic properties, whereas odd-numbered alkanes are anisotropic and pack less effectively next to one another (Yang et al., 2016). There is a possibility that the odd-even effect may have resulted in the differences in physicochemical properties such as molecular shape or packing efficiency of the odd chain *n*-nonane molecules, which could cause cell membrane entrapment or impact molecular interactions with the BEHP solvent, reducing *n*-nonane's partitioning into the aqueous phase compared to *n*-decane (Kim et al., 2007, Dhiman et al., 2022). Chapter 4.5 sets out to explore the impact of substrate solubility and partitioning across the LL interface.

Another striking observation in BR2 was that the TOF achieved by CYP153A6 whole cells on *n*-octane decreased from 24 to 48 hours by 1.3-fold while increasing for *n*-decane by 1.6-fold (Table 4.6). This observation is explained by the lower toxicity of *n*-decane compared to *n*-octane. Compound toxicity is indicated by logarithmic ratio of the partition coefficient ($\log P_{OW}$) in relation to the standard 1:1 octanol to water. As highlighted by Grant et al. (2011), longer-chain alkanes are less toxic to whole cells due to their higher $\log P_{OW}$ values, compared to 1-octanol with a $\log P_{OW}$ of 2.92 which cause cellular membrane disintegration (Table 4.7) (Olaofe et al., 2013, Grant et al., 2011). Therefore, the increase in the TOF achieved on *n*-decane indicated by its higher $\log P_{OW}$ value of 5.98 compared to 4.93 for *n*-octane (Table 4.7). Therefore, by 48 hours, whole cells show more instability and deactivation in the presence of *n*-octane than *n*-decane. The decrease in TOF from 24 to 28 hours on *n*-octane was attributed to an increase in 1-octanol concentration, resulting in toxicity to the whole cells. This result was supported findings by Olaofe et al. (2013) who demonstrated that product toxicity occurred with an increase in 1-octanol concentration. Time course data analyses and the aspects of compound toxicity throughout time is further illustrated and discussed in Chapter 4.2.3.

Table 4.7 The $logP_{OW}$ values of alkanes and their primary alcohol products (Tsukagoshi and Aono, 2000, Rojas et al., 2004)

Alkane	**$logP_{ow}$**	**Primary alcohol**	**$logP_{ow}$**
n-octane	4.93	1-octanol	2.92
n-nonane	5.45	1-nonanol	3.40
n-decane	5.98	1-decanol	4.00

The difference in TOFs achieved between alkanes of different lengths is greater than the relatively smaller differences observed between different substrates by purified enzymes. The impact of transport of substrate across the cellular membrane on the decrease of the TOF achieved on *n*-nonane to a greater extent than in systems using purified enzymes is demonstrated in Bordeaux et al. (2011). For example, Bordeaux et al. (2011) the TOF achieved by whole cells at 25˚C on *n*-octane was 5.5-fold larger than towards *n*-decane, whereas for purified CYP153A13 this difference was only 2-fold (Tables 4.8). In purified enzyme 2LPSs, differences in TOFs achieved on different alkane chain lengths can result from interfacial LL transport, substrate-solvent molecular interactions, substrate solubility, or affinity between the enzyme and substrate. In comparison, whole cell systems experience an additional cellular membrane transport barrier (White, 2022). Overall, the results illustrated that alkane transport across the cellular membrane or LL interface could be a rate-limiting step with an increase in alkane chain length and decrease in solubility. This supported previous authors Grant et al. (2014) and White (2022) who showed that *n*-alkane transport across the cellular membrane and LL interface was the rate-limiting step in whole cell *E. coli* bioconversion of C_7-C_{16} alkanes.

Table 4.8 Biohydroxylation parameters of CYP153A13 whole cells on different alkane chain lengths determined experimentally in the present study with comparison to literature studies

	This study (BR1)[1]		Pennec et al. (2015)		Bordeaux et al. (2014)		Bordeaux et al. (2011)		Bordeaux et al. (2011)	
Biocatalyst	*E. coli* CYP153A13		*E. coli* CYP153A13		Purified CYP153A13 from HCDC		Purified CYP153A13		*E. coli* CYP153A13	
Growth medium	LB		ZYP-5051 autoinduction		2xYT or LB or TB		LB or TB		LB or TB	
Induction	IPTG for 24 hrs		48 hrs		IPTG for 24 hrs		IPTG for 25 hrs		IPTG for 25 hrs	
Biotransformation medium	2LPS		2LPS		2LPS	2LPS	Monophasic		2LPS	
Buffer	200 mM sodium phosphate pH 7.2		200 mM Tris-HCl pH 8		100 mM potassium phosphate pH 7.4	100 mM potassium phosphate pH 7.4	100 mM potassium phosphate pH 7.4		100 mM potassium phosphate pH 7.4	
Time of measurement (hrs)	24		8		0.25	78	0.17		2	
Biotransformation temperature (˚C)	20		20		25		25		25	
Biomass concentration ($g_{DCW} \cdot L_{BRM}^{-1}$)	5.8		5		n.d		n.d		7.5	
CYP concentration ($\mu mol \cdot L_{BRM}^{-1}$)	0.23		4.3		0.5		0.15		n.d	
Substrate	C_8	C_9	C_8	C_{10}	C_8		C_8	C_{10}	C_8	C_{10}
TOF ($\mu mol_{1-alkanol} \cdot \mu mol_{P450}^{-1} \cdot min^{-1}$)	3.8	2.7	1.7	0.31	16	0.7	57	30	53	9.7

n.d= not determined

1: Please note that BR1 included tests on the activity of CYP153A6 and CYP153A13 whole cells on C_8 and C_9 from 0 to 72 hours. BR2 included tests on the activity of CYP153A6 and CYP153A13 whole cells on C_8, C_9 and C_{10} from 24 to 72 hours. BR3 included tests on the activity of CYP153A6 and CYP153A13 whole cells on C_8, C_9 and C_{10} from 0 to 12 hours.

The *in vivo* CYP153A13 TOF achieved for *n*-octane (3.8 $\mu mol_{1-octanol} \cdot \mu mol_{P450}^{-1} \cdot min^{-1}$) was 1.4-fold higher than on *n*-nonane (2.7 $\mu mol_{1-nonanol} \cdot \mu mol_{P450}^{-1} \cdot min^{-1}$) (Table 4.8). These results reflected the previous evidence that reactions of shorter chain alkanes exhibit higher TOFs, which can be attributed to their higher solubility and availability in the aqueous phase facilitating higher frequency of contact between the enzyme and substrate. This rationale strongly underpins the results in this chapter, but exceptions were noted, such as those related to the odd-even effect of liquid alkanes. The present study set out to explore further the effect of alkane chain length and to provide understanding of this effect. The findings presented here comparing the C_8 and C_{10} even-length alkanes through demonstrating the impact of substrate solubility on the TOFs achieved by whole cell biocatalysts supported findings on the 5-fold higher TOFs achieved on *n*-octane compared to *n*-decane by Bordeaux et al. (2011) and Pennec et al. (2015) (Table 4.8). Further, this study addresses the knowledge gap regarding CYP153's activity on *n*-nonane, a substrate that appears to have received little attention in previous studies.

As discussed, when comparing the activity of CYP153A6 to previous studies, the variability of TOFs achieved by CYP153A13 in different studies was attributed due to the use of different process, reaction, and induction conditions (Table 4.8). The TOF achieved by CYP153A13 whole cells at 24 hours was 3.8 $\mu mol_{1-octanol} \cdot \mu mol_{P450}^{-1} \cdot min^{-1}$, while Pennec et al. (2015) reported a TOF of approximately half at 1.7 $\mu mol_{1-ocanol} \cdot \mu mol_{P450}^{-1} \cdot min^{-1}$ at 8 hours (Table 4.8). The pH, buffer and biomass concentration used in the present study differed from those in Pennec et al. (2015)'s study (Table 4.8). Additionally, the present study utilised IPTG induction to express the CYPs, whereas Pennec et al. (2015) used an autoinduction medium (Table 4.8).

The importance of reporting reaction conditions accurately is highlighted by the observed variability observed in the results obtained by different authors investigating CYP153A6 and CYP153A13. Researchers can prevent the reaching of misguided conclusions as well ensure reproducibility and reliability of their results by providing detailed information regarding experimental setup and reaction conditions. These findings supported previous discussions by several authors and members of the European Federation of Biotechnology Section on Applied Catalysis on the practical guidelines of reporting experiments according to the Standards for Reporting Enzymology Data (Gardossi et al., 2010, Halling et al., 2018, Tipton et al., 2014, Swainston et al., 2018). Furthermore, the alkane activation research group, part of the c*Change Centre of Excellence in Catalysis Research in South Africa, set out a standardised reporting method to enable the comparison of catalysts across the biological, chemical engineering, and chemistry fields, which was applied throughout the current study (Appendix 7.6).

In agreement with the results of BR1, BR2 also showed that after 48 hours, the CYP153A13 whole cell catalyst achieved a TOF that was 1.4 times higher on *n*-octane than *n*-nonane (Table 4.9). CYP153A13 whole cells showed no *in vivo* activity on *n*-decane in BR2 and BR3, which demonstrated a lack of frequent contact between CYP153A13 and substrate because of a combination of two factors: (1) low concentrations of full-length correctly folded CYP153A13 and (2) low availability of *n*-decane in the aqueous phase and the cytoplasm of the cell. In contrast to these results, Pennec et al. (2015) obtained a TOF on decane of 0.31 $\mu mol_{1-decanol} \cdot \mu mol_{P450}^{-1} \cdot min^{-1}$, obtaining productive CYP153A13 cultures after 8 hours of reaction at 20˚C using an autoinduction medium for 48 hours. This may have been due to the variability seen in the activity of CYP153A13 cultures. In fact, Bordeaux et al. (2014) noted that some CYP153A13 cultures were unproductive at 25˚C, which was too the case in the present study, in part because of lower CYP153A13 concentrations and in part because of the low *n*-decane solubility.

Table 4.9 Biohydroxylation performance of CYP153A13 on n-octane, n-nonane, and n-decane (BR2)

Biocatalyst	*E. coli* CYP153A13					
Growth medium	LB					
Induction time (hrs)	24					
Biotransformation medium	2LPS consisting of organic (alkane and BEHP) and aqueous medium					
Biotransformation temperature (˚C)	20					
Biomass concentration ($g_{DCW} \cdot L_{BRM}^{-1}$)	5.4					
CYP concentration ($\mu mol \cdot L_{BRM}^{-1}$)	0.175					
Time of measurement (hrs)	24			48		
Alkane	Octane	Nonane	Decane	Octane	Nonane	Decane
TOF ($\mu mol_{1-alkanol} \cdot \mu mol_{P450}^{-1} \cdot min^{-1}$)	0.21	0	0	0.16 ± 0.0013	0.12 ± 0.012	0

The challenges of the aqueous solubility of hydrophobic substrates in whole cell biocatalytic 2LPSs were explored and the substrate scope of CYP153A6 and CYP153A13. It was illustrated that biohydroxylation was successful for CYP153A6 whole cells on *n*-octane, *n*-nonane, and *n*-decane. CYP153A13 whole cells catalysed the primary hydroxylation of *n*-octane and *n*-nonane. These results represent the apparent activity and substrate scope of CYP153A6 and CYP153A13 whole cell biocatalysts for the hydroxylation of medium chain length alkanes. These single time-point analyses of *in vivo* hydroxylation assays enabled easy comparison with the results of previous studies and offered a deeper understanding of the impact of solubility on alkane transport across the membrane and LL barriers. Therefore, while single time-point analyses were beneficial, the present study set out to investigate the factors that impact reaction rates and kinetics over time. These are addressed in Chapter 4.2.3 which investigates the impact of solubility on TOFs over time, while Chapter 4.3 investigates the impact of solubility on the kinetics of whole cell CYP153A6 and CYP153A13.

Longer chain alkanes are less soluble than shorter chain lengths, which limits the frequency of ES contact due to transport barriers across the LL interface and the cell membrane. Compelling evidence was presented that aqueous alkane solubility and chain length play a role in the extent of this transport, affecting the resultant TOFs achieved by whole cell resting cell biocatalyst cultures. Interestingly, the lower achieved TOFs on *n*-nonane compared to *n*-decane suggest that the odd-even effect of liquid alkanes had a greater impact than alkane solubility, in the case of *n*-nonane and *n*-decane. These results suggested that alkane transport across either the membrane of LL interfacial barriers was found to be affected by whether the alkane chain length was an even or odd number.

4.2.2 Comparison of biohydroxylation activity of multicomponent to fusion CYP153 enzyme whole cell biocatalysts on medium chain length alkanes

Biohydroxylation reactions catalysed by CYP153A13 whole cells achieved lower TOFs at 24 hours than by CYP153A6 on all the medium-chain alkanes investigated in this study due to previously discussed low expression levels of CYP153A13 (Tables 4.5, 4.6, 4.8, 4.9; Figure 4.7). In BR1, the TOF achieved by CYP153A6 whole cell biohydroxylation on *n*-octane was four-fold higher than by CYP153A13 at 24 hours (Figure 4.8). Similarly, the TOF achieved by CYP153A6 whole cells on *n*-nonane was three-fold higher than that achieved by CYP153A13 at 24 hours. The results correlated with BR2, showing that the TOF achieved by CYP153A6 whole cells on *n*-octane was five-fold higher than that of CYP153A13 at 48 hours and two-fold higher on *n*-nonane (Figure 4.8). These results support findings by Pennec et al. (2015), who reported a five-fold larger TOF achieved by CYP153A6 whole cells compared to CYP153A13 on *n*-octane.

Although fusion enzymes possess superior coupling efficiencies, these benefits are outweighed by low enzyme expression levels. Fusion enzymes, comprising a single unit of the P450 component fused to the electron transfer proteins, are often too large and complex to express in hosts such as *E. coli*, which results in low expression of full-length, fully functional enzymes (LaVallie and McCoy, 1995). CYP153A6 expressed as a full-length and correctly folded enzyme in higher concentrations, increasing its active site availability to catalyse the hydroxylation reaction compared to CYP153A13. This complicates determining whether fusion enzymes and their superior coupling efficiencies are superior to multicomponent enzyme systems. The expression strategy of CYP153A13 should be improved before concluding the benefit of superior coupling efficiencies of fusion enzymes compared to multicomponent enzyme systems. However, since TOF however normalises by the production of product per a unit of expressed enzyme and is measured by $\mu mol_{1-alkanol} \cdot \mu mol_{P450}^{-1} \cdot min^{-1}$, the low expression levels may not be the largest contributing factor to the lower TOFs achieved by CYP153A13 compared to CYP153A6. Another explanation for the results showing that CYP153A6 achieved superior TOFs in comparison to CYP153A13 may be that *n*-octane and *n*-nonane had better affinities with CYP153A6 than CYP153A13.

In summary, the results above are based on TOF at a single sampling point of the reaction and were used to derive general implications. The analyses should be interpreted cautiously because multiple factors contribute to a system's productivity and apparent activity (Figure 4.8). Several factors intervene with the measured TOFs, including the enzyme expression levels, electron coupling efficiency, substrate solubility, substrate-solvent interactions, and the affinity between substrate and enzyme to form an ES complex within the cytoplasm of a whole cell biocatalyst (Figure 4.8).

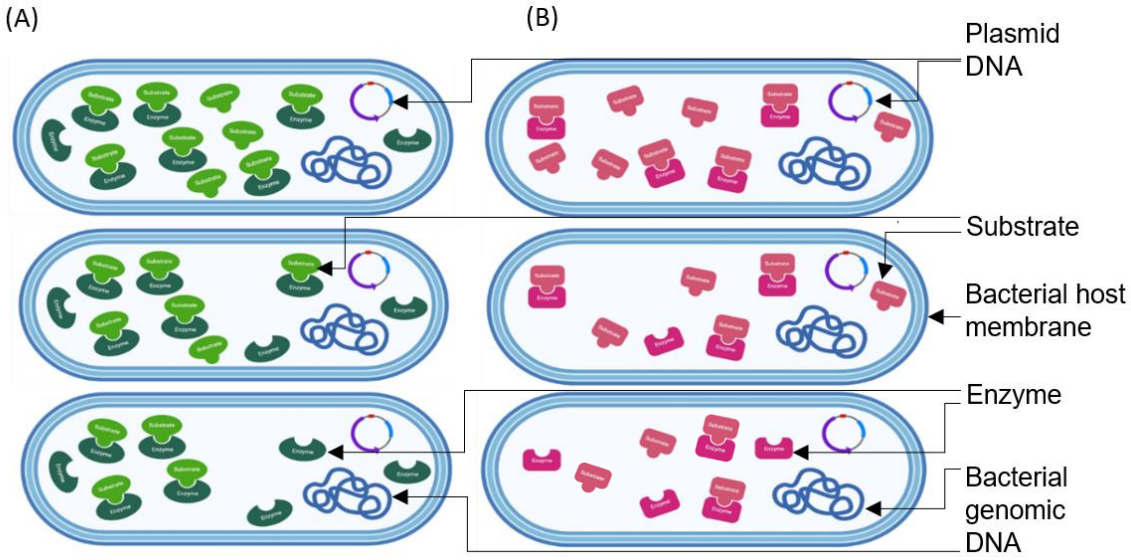


Figure 4.8 A diagram illustrating the roles of substrate solubility and enzyme expression in enzyme-substrate contact. The top row shows whole cells containing *n*-octane, the middle *n*-nonane, and the bottom row *n*-decane, illustrating that a decrease in substrate solubility results in lower substrate amounts in the cytoplasm of the whole cells. Column (A) shows whole cells expressing CYP153A6, and column (B) shows whole cells expressing CYP153A13 in a lower concentration. Therefore, column B illustrated that the lower concentration of CYP153A13 resulted in fewer active sites being available to contact the substrate compared to whole cells in column A which showed a higher concentration of CYP153A6.

4.2.3 Analyses of turnover frequency of the bioconversion of alkanes into primary alcohols catalysed by CYP153 whole cells over time

The analyses of an entire reaction progress curve can provide more insight into the components of a whole cell biocatalytic 2LPS over time, such as product accumulation, substrate depletion, or cellular health. The reaction rates in terms of TOF as a function of time were evaluated from 3 to 72 hours to understand factors that affect the turnover of alkane substrate to primary alcohol product of each ES combination (Figure 4.9). In BR1 (Figure 4.9A), the data points represent a single sample, while in BR2 and BR3 (Figure 4.9B and C), the data points represent an average of two samples. The profile of TOF as a function of time possesses characteristics such as the rate of increase or decrease in TOF over time, plateaus, and maxima. These characteristics were examined and compared to identify the limiting parameters of the medium chain length alkane primary hydroxylation reactions catalysed by CYP153A6 and CYP153A13, respectively (Figures 4.9A to C).

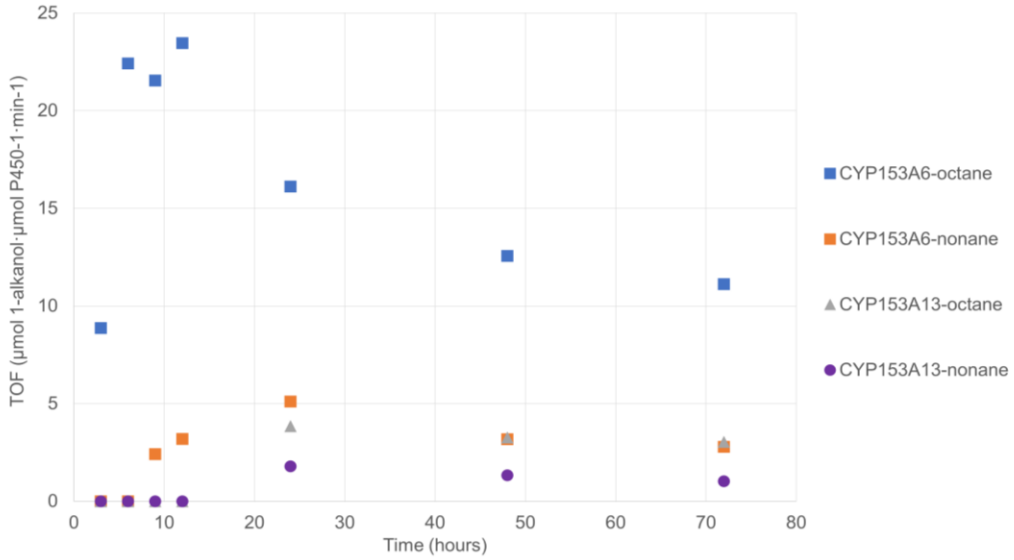
The TOF maxima achieved were determined directly from the plots in Figure 4.9. In BR1, CYP153A6 whole cells achieved a maximum TOF on *n*-octane (23.5 $\mu mol_{1-octanol} \cdot \mu mol_{P450}^{-1} \cdot min^{-1}$) at 12 hours, in half the time and at six times the magnitude compared to the maximum TOF achieved by CYP153A13 whole cells (3.84 $\mu mol_{1-octanol} \cdot \mu mol_{P450}^{-1} \cdot min^{-1}$), for reasons discussed in the next paragraph (Figure 4.9A). The TOF of the CYP153A6 whole cells decreased to 16.12 $\mu mol_{1-octanol} \cdot \mu mol_{P450}^{-1} \cdot min^{-1}$ by 24 hours. For *n*-nonane, CYP153A6 achieved a three-fold higher TOF maximum of 5.10 $\mu mol_{1-nonanol} \cdot \mu mol_{P450}^{-1} \cdot min^{-1}$ than the TOF maximum achieved by CYP153A13 of 1.79 $\mu mol_{1-nonanol} \cdot \mu mol_{P450}^{-1} \cdot min^{-1}$ at 24 hours. After 24 hours, the TOF of *n*-nonane by CYP153A6 decreased more rapidly than CYP153A13 (Figure 4.9A). In BR2, the maximum TOF achieved by CYP153A6 whole cells on *n*-nonane (0.35 $\mu mol_{1-nonanol} \cdot \mu mol_{P450}^{-1} \cdot min^{-1}$) was two-fold higher than the TOF achieved by CYP153A13 whole cells (0.16 $\mu mol_{1-nonanol} \cdot \mu mol_{P450}^{-1} \cdot min^{-1}$) at 72 hours. It was noted that the TOF by CYP153A6 lay in the range 0.28 to 0.35 $\mu mol_{1-nonanol} \cdot \mu mol_{P450}^{-1} \cdot min^{-1}$between 24 and 72 h, whereas the TOF of

CYP153A13 lay in the range 0.12 to 0.16 $\mu mol_{1-nonanol} \cdot \mu mol_{P450}^{-1} \cdot min^{-1}$ between 48 and 72 h. This indicated that a plateau had been reached by CYP153A6 and CYP153A13, and that the plateau of TOF reached by CYP153A6 was approximately 2-fold higher than that of CYP153A13.

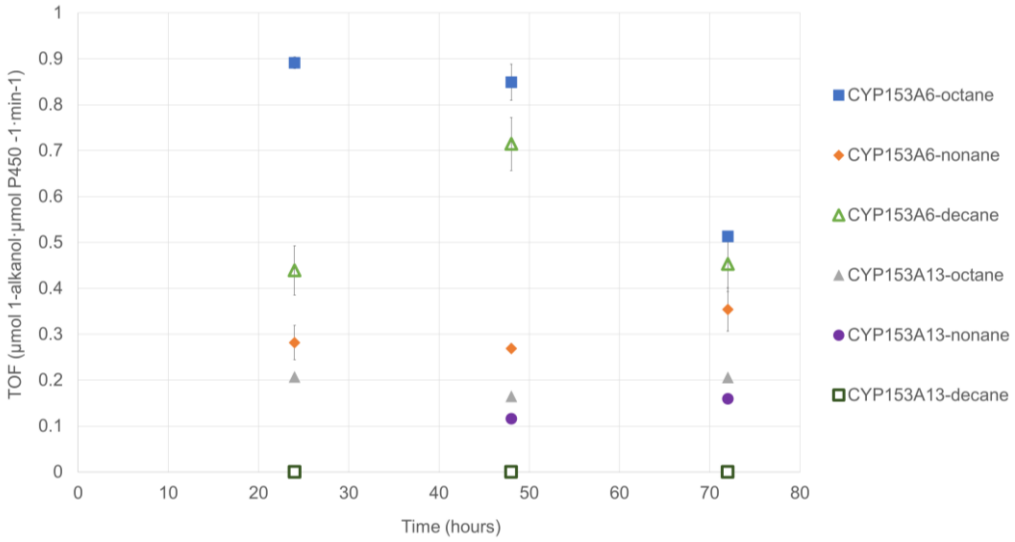
There are a few reasons that may be contributing to these results. Firstly, plateaus of TOF reached over time most likely signifies the phase of the reaction in which the enzyme is performing at its maximum capacity. Secondly, CYP153A6 may have a greater affinity for *n*-octane and *n*-nonane and a superior catalytic mechanism with these substrates compared to CYP153A13, and/or it could be related to the higher concentration of CYP153A6. Compared to CYP153A6, these analyses highlighted that low expression levels of CYP153A13 may have decreased the achievable TOF maximum. Additionally, the TOF maxima achieved by CYP153A6 whole cells were reached more rapidly than by CYP153A13 whole cells, illustrating that CYP153A6 displayed faster rates of substrate turnover into the product due to its higher cytoplasmic concentrations. Low CYP153A13 expression levels translated into less frequent interaction between the enzyme's active site and substrate to form the ES complex, also known as the reaction's transition state.

Similarly demonstrated by results in BR2, showing a three-fold higher maximum TOF of 0.89 $\mu mol_{1-octanol} \cdot \mu mol_{P450}^{-1} \cdot min^{-1}$ by CYP153A6 compared to CYP153A13 whole cells with a maximum TOF of 0.21 $\mu mol_{1-octanol} \cdot \mu mol_{P450}^{-1} \cdot min^{-1}$ at 24 hours (Figure 4.9B). The TOF as a function of time for BR3 demonstrated that the catalytic conversion of *n*-octane by CYP153A6 whole cells reached a maximum TOF (3.95 $\mu mol_{1-octanol} \cdot \mu mol_{P450}^{-1} \cdot min^{-1}$) at 12 hours, at three times the magnitude of the maximum TOF of 1.56 $\mu mol_{1-octanol} \cdot \mu mol_{P450}^{-1} \cdot min^{-1}$ by CYP153A13 whole cells at 9 hours (Figure 4.9C). The TOF maximum of CYP153A13 on *n*-octane was reached at 3 hours before CYP153A6, but this was because the TOFs achieved by CYP153A13 levelled off from 6 to 12 hours, whereas CYP153A6 showed a general trend of increase over time (Figure 4.9C). The characteristic of CYP153A6 showing rapid increases and decreases in TOF over time and CYP153A13 showing stable plateaus in TOF over time was similarly observed in BR1 and BR2 (Figures 4.9A and B). These observations highlighted the importance of interpreting biocatalytic activity over time rather than single-time point time analyses.

(A)



(B)



(C)

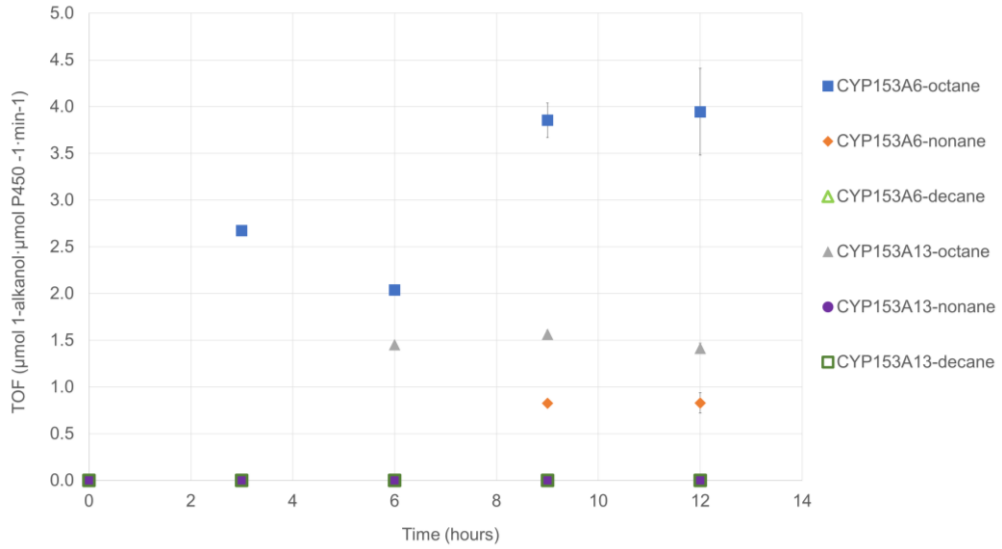


Figure 4.9 The TOF as a function of time plot of CYP153A6 and CYP153A13 whole cells catalysing the primary hydroxylation of longer chain alkanes at 20˚C, preceding 24 hours IPTG induced enzyme expression in (A) biological repeat 1 (B) biological repeat 2 and (C) biological repeat 3.

In all three biological repeats, the impacts of substrate solubility on the profiles of TOFs achieved over time was illustrated. For example, reactions on *n*-octane showed the most rapid increase to reach the TOF maxima and the most rapid decrease after the maxima by both whole cell biocatalysts, while the biohydroxylation on *n*-nonane displayed the most gradual change in TOF over time (Figure 4.9). This can be explained by the higher solubility of *n*-octane, resulting in higher cytosolic concentrations than *n*-nonane, leading to more frequent contact between enzyme and substrate within a unit of time, resulting in a higher achievable TOF maxima (Figure 4.9). The cytoplasmic concentration of alkane is dependent on its transport across both the transmembrane and LL interface. Transport across the LL interface is related to the partitioning behaviour of substrate at phase equilibrium, which impacts the substrate availability to the whole cell biocatalyst in the aqueous phase, investigated further in Chapter 4.5 (Domańska et al., 2015, Kim et al., 2007).

The decline in TOFs over time by whole cell biocatalysts, clearly observed in Figure 4.9A, may be caused by several variables. For instance, declining cellular health, enzyme deactivation, enzyme inhibition by product or substrate depletion could all be factors contributing to the decline of the TOF over time (Figures 4.9A to C). In fact, the impact of the relative toxicity of the different alkane substrates or primary alcohol products on reaction rates was demonstrated (Table 4.9). As mentioned in single time-point analyses in Chapter 4.2.1 (Table 4.7), compounds with a lower range $\log P_{OW}$ values (from one to four) disintegrate *E. coli's* cellular membrane (Grant et al., 2011) As alkane chain length increases, the toxicity toward the cell decreases.

The $\log P_{OW}$ of 1-octanol is 2.92 making it the most toxic compared to the other studied substrates and products. Interestingly, the analyses of TOFs as a function of time in Figure 4.9 illustrated a trade-off between substrate availability and toxicity. The TOF by CYP153A6 for the bioconversion of *n*-octane showed a large decrease from 24 to 48 hours (Figure 4.9B). In contrast, the TOF of CYP153A6 bioconversion of *n*-decane to 1-decanol increased from 24 to 48 hours. The toxicity of *n*-decane and 1-decanol appeared to only impact cellular health after 48 hours, characterised by a decrease from 48 to 72 hours (Figure 4.10B).

Clearly, shorter chain lengths compromised cellular health more quickly than longer chains. Octane is more soluble in the aqueous phase than *n*-decane, resulting in a higher TOF to produce 1-octanol, resulting in higher toxicity levels to cellular health in a shorter timespan (Figure 4.9B). Shorter chain alkanes and alcohols are more soluble in the aqueous phase and therefore more available, facilitating a higher rate of increase in TOF over time than longer chain lengths, yet show higher toxicity towards whole cells due to their lower $\log P_{OW}$ values. These results support evidence by Olaofe et al. (2013), who showed that solvent toxicity negatively impacted cellular health, limiting 2LPS whole cell biocatalysis productivity (Olaofe et al., 2013). Furthermore, the results also supported a study by Grant et al. (2011) who showed that higher volumetric rates were obtained for converting *n*-dodecane to 1-dodecanol by *E. coli* expressing the alkB complex compared to the bioconversion of *n*-octane, attributed to the lower toxicity of longer-chain alkanes and their primary alcohol products.

These findings justified the application of cosolvent BEHP to alleviate substrate and product toxicity by regulating the concentrations of these compounds in the biocatalyst-residing aqueous phase (Lewandowski, 2008, Olaofe et al., 2013, Kim et al., 2007). Moreover, the depletion of nutrients would have also contributed to the decline in TOFs achieved towards 72 hours, which could be improved by using an intermittent feeding system (Figure 4.9) (Olaofe et al., 2013). In CYP153A6 whole cell 2LPSs, implementing strategies to maintain cellular health over the reaction period in the presence of toxic products while adjusting phase equilibrium to increase substrate availability in the aqueous phase would

increase the TOFs achieved throughout the 72-hour reaction period. On the other hand, for CYP153A13 whole cell 2LPSs, it was evident that increased protein expression, enzyme concentration, and ES contact needed improvement to increase the achieved TOFs. Since CYP153A6 showed higher expression levels and concentrations, these systems already experienced a higher frequency of ES contact.

Overall, the comparison of TOF maxima between whole cell biocatalysts and substrates led to similar observations as those in single-time point analyses in Chapter 4.2.1 and 4.2.2, which enabled comparison to previous studies that reported the TOF at 24 hours, regardless of when the maximum TOF of the reaction was reached. In contrast, the analyses here placed the reaction's performance into the context of its progress over time (Figure 4.9) Although true ES compatibility was not revealed in the present chapter, the analyses provided approximate suitability and overall functionality of CYP153 whole cell biocatalysts for the hydroxylation of medium-chain alkane substrates in a 2LPS. In an ideal system, the profile of reaction progress plots (such as TOF as a function of time) is solely influenced by intrinsic kinetics, including the substrate turnover to product, which is primarily dictated by the affinity of enzyme towards substrate, which can be indicated by what is referred to as the Michaelis-Menten constant (K_m). Therefore, in an ideal system, there is no influence of substrate-solvent interactions or limitations arising from substrate solubility, LL interfacial transport, transmembrane transport, microbial health, and the extent of enzyme expression levels on the TOF outcome exhibited by a whole cell biocatalyst.

4.3 Determination of apparent Michaelis-Menten kinetic parameters for CYP153 whole cell biocatalytic primary hydroxylation of medium-chain alkanes in a 2-Liquid-Phase-System

Turnover frequency (TOF) as a function of time analyses provided information about the rate of product formed per unit of time per amount of enzyme present, but these plots do not provide information on the essential properties of an enzyme under substrate saturation. At substrate saturation, the substrate occupies all active sites of the enzyme, and this signifies the maximum velocity of a reaction. The maximum reaction velocity of the respective enzyme-substrate (ES) pairs was determined using the kinetic Michaelis-Menten (MM) model (Equation 4.1) model. The kinetics of whole cell biocatalytic medium chain alkane primary hydroxylation was evaluated using the MM model to establish the relationship between alkane concentration and alkane hydroxylation rate. Enzyme saturation kinetics were described to develop deeper insight into how easily the CYP153A6 and CYP153A13 enzymes are saturated with a particular alkane substrate, respectively defined as the MM constant (K_m). The maximum rate of production at the saturation point (v_{max}) of each ES pair was also determined.

$$v = \frac{d[P]}{dt} = \frac{-d[S]}{dt} = \frac{v_{max}[S]}{K_m^c + [S]} \qquad \textit{Equation 4.1}$$

This kinetic description allowed a deeper understanding of the substrate scope of CYP153A6 and CYP153A13 and their respective reaction mechanisms. The formation of enzyme-substrate (ES) complex is dependent on both ES interactions and substrate-solvent interactions, which is indicated by the apparent MM constant K_m^c. The term 'apparent' kinetics appertains to the determination of parameters that include variables such as substrate-solvent interactions, substrate transport into the whole cells, and even substrate concentrations beyond the substrate's solubility limit i.e., excess substrate (Schrewe et al., 2014).

Classical enzyme kinetic studies typically construct the MM curve by plotting data points of several experiments determining the initial reaction velocities (v_0) starting with different initial substrate concentrations. Linear regression of the early stages of the resultant product concentration as a function of time is performed and the differential is used to obtain multiple initial reaction velocities v_0. The relationship between substrate concentration and v_0 is established by the hyperbolic MM Equation 4.1.

In 2-liquid-phase-systems (2LPSs), independently controlling the substrate concentration in the aqueous phase was challenging since it depends on thermodynamics and mass transfer across the liquid-liquid (LL) interface. This substantiated the need for the application of excess substrate concentration to increase ES contact. Since substrate was applied in excess, substrate utilisation rates were difficult to determine, and therefore, product formation rates were used instead. Furthermore, the reaction system was selected so that oxygen was not limiting to the kinetics of the reaction, using data from a previous study performed in our laboratory by Meissner et al. (2013) showing that oxygen transfer rates in both the large- and small-scale systems used had minimal effect on alkane oxidation using CYP153 whole cells. This study utilised the same small-scale system in which it was demonstrated that increased oxygen availability did not increase reaction rate

The concentration of product formed at selected time points was collected experimentally, whereas the MM model is formulated with respect to rates. Therefore, an expansion upon the classical enzyme kinetics method was required to enable MM parameter determination from a single product concentration versus time curve. As such, product concentration as a function of time data was collected experimentally using gas chromatography (GC) (Chapter 3.2.3) and fitted to the MM equation to

estimate the v_{max} and the K_m constant of each reaction using least squares nonlinear regression (NLR) (Chapter 3.4.2.1).

Experiments comprising of biological repeat 1 (BR1) with assays of both whole cell biocatalysts on *n*-octane and *n*-nonane from 0 to 72 hours and a biological repeat 2 (BR2) comprised of duplicates of both whole cell biocatalysts on *n*-octane, *n*-nonane, and *n*-decane from 24 to 72 hours were used in these analyses. In biological repeat 3 (BR3), both whole cell biocatalysts were tested in duplicate on *n*-octane, *n*-nonane, and *n*-decane in the initial phases of the reaction from 0 to 12 hours.

Each biological repeat fulfilled unique roles of kinetic analyses. Initial reaction velocities of the primary alkane hydroxylation reactions catalysed by CYP153 whole cells were calculated by differentiating time-course product concentration data within the initial stages of the reaction, where the product concentration was showing a linear increase with time. Initial reaction velocities are a useful indicator of enzyme productivity because the reaction is not yet impacted by enzyme deactivation or substrate depletion (Figure 4.10). Initial reaction velocities were determined by estimating the instantaneous rate by determining the gradient of the tangent to the linear section of the progress reaction curve (Srinivasan, 2022).

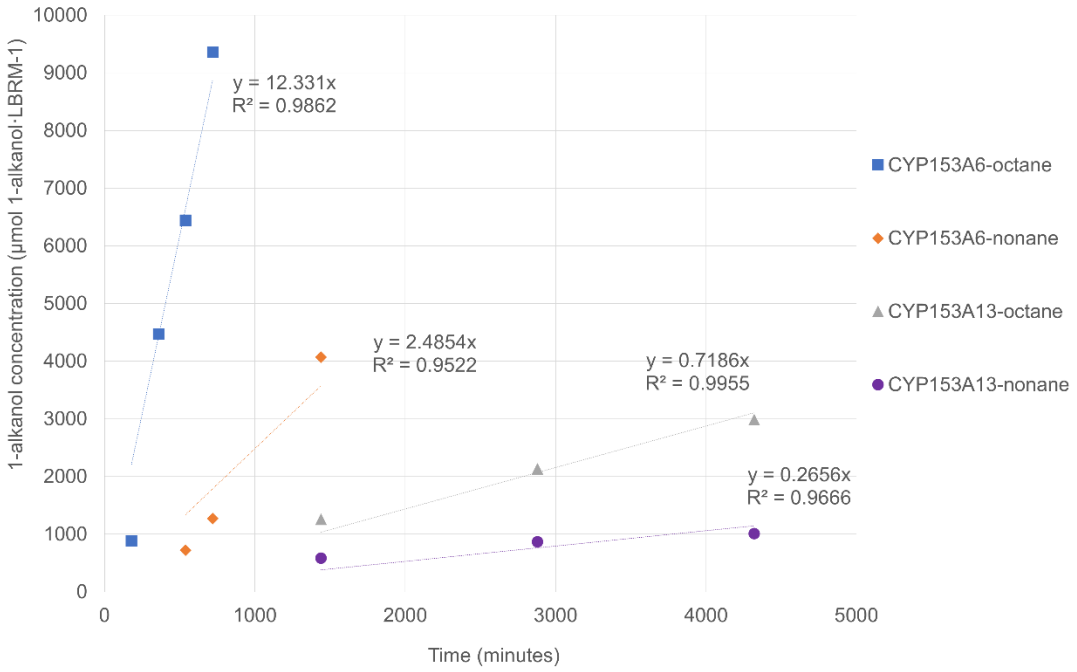
CYP153A6 and CYP153A13 recombinant cultures were grown, and enzyme expression was induced using IPTG at mid-exponential phase (OD_{600} of 0.8 to 1) and allowed to incubate at 20˚C for 24 hours at 200 rpm. Following this, cultures were resuspended at a concentration of 3-5 $g_{DCW} \cdot L_{BRM}^{-1}$ in buffer to a total volume of 1 mL. A volume of 200 µL of alkane dissolved in 100 µL cosolvent BEHP was then applied to initiate the reaction (Chapter 3.1.5.1).

The concentration of alkane placed in the 2LPS was determined within the working volume (WV), which excludes the volume of the cosolvent (Table 4.10). The initial substrate concentrations in the WV in the 2LPSs are specified in Table 4.10. Amber vials were sacrificed for product extraction at selected time points to measure the product concentration using gas chromatography (GC) (Chapter 3.2.3).

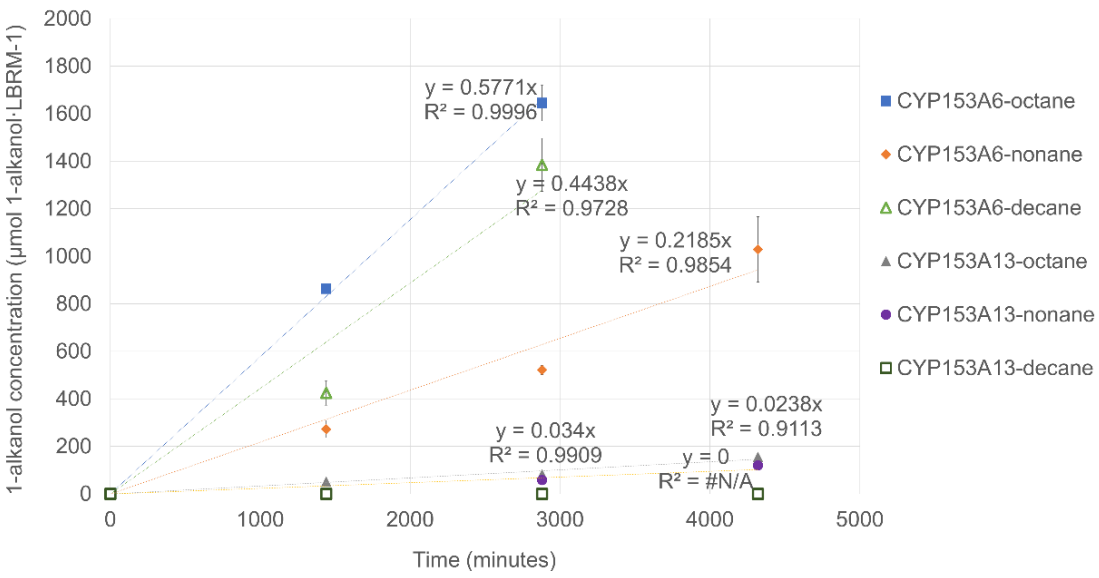
Table 4.10 Initial concentrations of alkane substrate in the working volume of the respective 2LPSs

Substrate	Amount of substrate in 2LPS (μmol)	Substrate concentration in working volume of 2LPS (μM)
n-octane	1231	1.03×10^6
n-nonane	1120	9.33×10^5
n-decane	1026	8.55×10^5

(A)


10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
1-alkanol concentration (µmol 1-alkanol·LBRM-1)
0
1000
2000
3000
4000
5000
Time (minutes)
y = 12.331x
R² = 0.9862
y = 2.4854x
R² = 0.9522
y = 0.7186x
R² = 0.9955
y = 0.2656x
R² = 0.9666
CYP153A6-octane
CYP153A6-nonane
CYP153A13-octane
CYP153A13-nonane

(B)


2000
1800
1600
1400
1200
1000
800
600
400
200
0
1-alkanol concentration (µmol 1-alkanol·LBRM-1)
0
1000
2000
3000
4000
5000
Time (minutes)
y = 0.5771x
R² = 0.9996
y = 0.4438x
R² = 0.9728
y = 0.2185x
R² = 0.9854
y = 0.0238x
R² = 0.9113
y = 0.034x
R² = 0.9909
y = 0
R² = #N/A
CYP153A6-octane
CYP153A6-nonane
CYP153A6-decane
CYP153A13-octane
CYP153A13-nonane
CYP153A13-decane

(C)

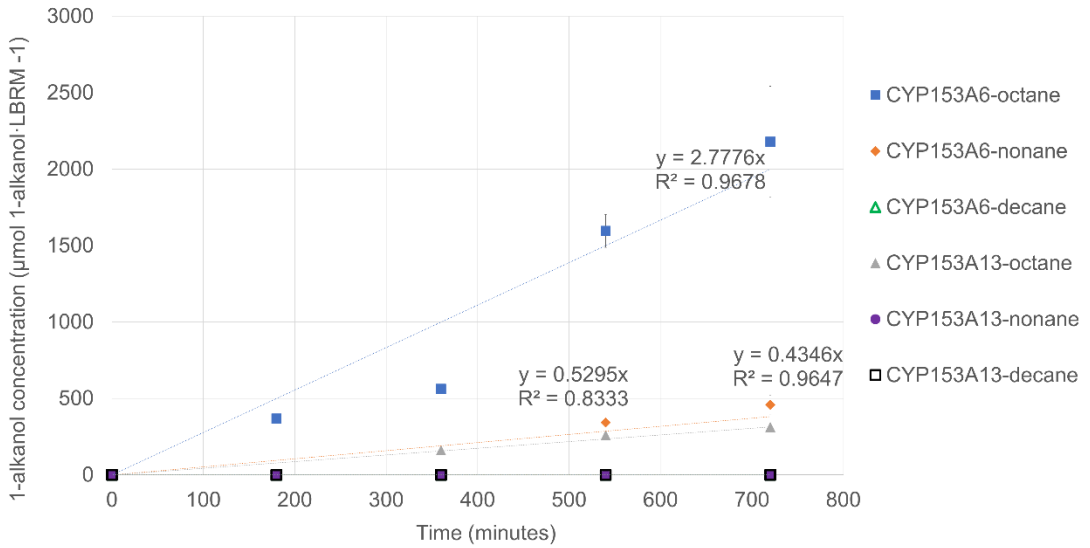


Figure 4.10 Graphs of biological repeats (A) BR1, (B) BR2 and (C) BR3 illustrating the measured 1-alkanol concentration at several time points and linear line of best fit to determine the gradient values to determine initial reaction velocities.

In BR1, the initial reaction velocities determined for each ES pair were from fastest to slowest: CYP153A6-octane> CYP153A6-nonane> CYP153A13-octane> CYP153A13-nonane (Figure 4.10A). In BR2, the fastest to the slowest initial reaction velocities for the ES pairs were ordered: CYP153A6-octane> CYP153A6-decane> CYP153A6-nonane> CYP153A13-octane> CYP153A13-nonane (Figure 4.10B). In BR3, the fastest to the slowest initial reaction velocities for the ES pairs were ordered: CYP153A6-octane> CYP153A6-nonane> CYP153A13-octane (Figure 4.10C). *In vivo* assays revealed that the initial reaction velocities achieved by CYP153A6 were consistently faster than reactions utilising CYP153A13 whole cells on the range of alkanes tested due to CYP153A13's low expression levels, as shown in previous sections. More specifically, the initial reaction velocity of the hydroxylation of *n*-octane biocatalysed by CYP153A6 whole cells was determined to be 12.33 $\mu mol_{1-octanol} \cdot L_{BRM}^{-1}$, which was 17-fold faster than the v_0 for reactions utilising CYP153A13 (0.719 $\mu mol_{1-octanol} \cdot L_{BRM}^{-1}$). Similarly, the v_0 of CYP153A6 whole cell hydroxylation on *n*-nonane (3.77 $\mu mol_{1-nonanol} \cdot L_{BRM}^{-1}$) was 15-fold faster than the biohydroxylation by CYP153A13 whole cells (0.266 $\mu mol_{1-nonanol} \cdot L_{BRM}^{-1}$).

Corresponding results were obtained in BR3 (Figure 4.10C). The initial reaction velocity of CYP153A6 whole cells on *n*-octane was seven-fold faster than CYP153A13 whole cells (2.96 versus 0.45 $\mu mol_{1-octanol} \cdot L_{BRM}^{-1}$). In BR2, CYP153A6 whole cell cultures achieved an initial reaction velocity of 0.577 $\mu mol_{1-octanol} \cdot L_{BRM}^{-1}$, which was 17-fold faster than the reaction utilising CYP153A13 whole cells on *n*-octane with an initial reaction velocity of 0.034 $\mu mol_{1-octanol} \cdot L_{BRM}^{-1}$. The results repeatedly demonstrated that the low expression of CYP153A13 in *E. coli* decreased the TOFs and initial reaction velocities achieved for alkane biohydroxylation, compared to CYP153A6. This was due to less frequent contact between enzyme and substrate due to lower CYP153A13 expression levels, decreasing the initial reaction velocity reaction rates.

The frequency of contact between enzyme and substrate can also decrease due to poor substrate solubility in the aqueous phase due to LL interfacial transport and substrate partitioning regulated by phase equilibrium. Longer alkane chain lengths further enhance this limitation of 2LPSs. The faster initial reaction velocities utilising CYP153 whole cells for the biohydroxylation of *n*-octane than *n*-nonane resulted from its higher solubility in the aqueous phase (Figure 4.10). For example, CYP153A13,

already expressed in low cytoplasmic concentrations, showed no *in vivo* activity on *n*-decane, which is longer in chain length and has a lower aqueous solubility than *n*-octane and *n*-nonane (Figure 4.10B).

These results contrasted with findings by Bordeaux et al. (2011), who reported that CYP153A13 whole cells had dissociation constants (K_D) of 29 and 16 μM for *n*-octane and *n*-decane, respectively and TOFs of 53 and 9.7 $\mu mol_{1-decanol} \cdot \mu mol_{P450}^{-1} \cdot min^{-1}$ at 25˚C instead of 20˚C. This higher temperature resulted in an increase in the solubility of *n*-decane into the aqueous phase, which possibly indicated that the reaction would have shown productivity in the present study, had the temperature increased by 5˚C. Although an increase in temperature would increase the solubility of *n*-decane, Olaofe et al. (2013) demonstrated that the specific production of 1-octanol by CYP153A6 whole cell biocatalysts at 25˚C reduced by 50% compared to 20˚C, which was associated with a faster rate of deactivation of CYP153A6 under 25˚C. The half-life of the biocatalyst decreased from 25 to 10 hours when the temperature increased from 20 to 25˚C and the half-life and specific 1-octanol concentration showed an even further decrease at 30˚C. This illustrates the importance of taking into account both substrate solubility and enzyme thermostability when optimising reaction conditions and achieving an optimal balance between these two parameters.

Another example that displayed the impact of substrate solubility and its partitioning into the aqueous phase was the faster initial reaction velocities obtained for *n*-octane compared to *n*-nonane catalysed by CYP153 whole cells in all three biological repeats (Figure 4.10).This observation was explicitly noted by the three-fold faster initial reaction velocity of CYP153A6 whole cell bioconversion of *n*-octane at 12.33 $\mu mol_{1-octanol} \cdot L_{BRM}^{-1} \cdot min^{-1}$ compared to the initial reaction velocity of 3.77 $\mu mol_{1-nonanol} \cdot L_{BRM}^{-1} \cdot min^{-1}$ on *n*-nonane (Figure 4.10A). Similarly, CYP153A13 whole cell bioconversion of *n*-octane had a three-fold faster initial reaction velocity of 0.719 $\mu mol_{1-octanol} \cdot L_{BRM}^{-1} \cdot min^{-1}$ than *n*-nonane with a v_0 of 0.266 $\mu mol_{1-nonanol} \cdot L_{BRM}^{-1} \cdot min^{-1}$ (Figure 4.10A). Also, in BR3, the initial reaction velocity of CYP153A6 for the bioconversion of *n*-octane was 2.97 $\mu mol_{1-octanol} \cdot L_{BRM}^{-1} \cdot min^{-1}$ which was six-fold faster than the bioconversion of *n*-nonane (0.53 $\mu mol_{1-nonanol} \cdot L_{BRM}^{-1} \cdot min^{-1}$)(Figure 4.10C). Octane is 3.3-fold more soluble in water than *n*-nonane and therefore has greater availability to the whole cell biocatalyst and to make contact with the active site of the CYP153 enzyme, facilitating faster reaction rates.

The aqueous solubility of the alkane in whole cell biocatalytic 2LPSs impacts the efficiency at which substrate is transported across the cellular membrane and LL interface, as mentioned in previous studies (Grant et al., 2011, van Tol et al., 1995, White et al., 2017), with lower solubility leading to delays in biocatalytic primary alcohol production. Together with Chapter 4.2's findings, these findings were supported by White et al. (2017), who identified alkane transmembrane transport as a major rate-limiting step in alkane whole cell hydroxylation. The most crucial aspect of biocatalytic conversion and velocity rates is the frequency of contact between the enzyme active site and substrate.

Interestingly CYP153A6 catalysed reaction of *n*-decane at an initial reaction velocity of 0.4438 $\mu mol_{1-decanol} \cdot L_{BRM}^{-1} \cdot min^{-1}$, two-fold faster than the initial reaction velocity of 0.2185 $\mu mol_{1-nonanol} \cdot L_{BRM}^{-1} \cdot min^{-1}$ on n-nonane (Figure 4.10B). This trend, also discussed in Chapter 4.2, was attributed to liquid alkanes' odd-even effect, which impacted transport across the cellular membrane and/or the LL interface. The number of carbons of an alkane impacts its molecular behaviour and physicochemical properties. Interestingly, these effects are related to the shape and nature of the alkane rather than its solubility. Overall alkane chain length, solubility, and shape - which differ in odd chains compared to even chains due to molecular asymmetry - have an impact on the rate of alkane whole cell

biohydroxylation reactions. These factors also have an impact on how the molecules interact with cellular membranes and cosolvent molecules.

4.3.1 Determination of the maximum velocity (v_{max}) and Michaelis-Menten constant (K_m^c) for whole cell alkane biohydroxylation reactions in a 2LPS using least squares nonlinear regression

The apparent kinetic parameters v_{max} and K_m^c were determined for the CYP153 whole cell biohydroxylation of medium-chain alkanes. The v_{max} describes the velocity of an enzyme-catalysed reaction under maximum enzyme functionality with all active sites saturated with substrate and, therefore, can assist in selecting the substrate concentration needed to obtain a maximal reaction velocity (Tipton et al., 2014). The MM constant K_m^c indicates the substrate concentration at which half the v_{max} is reached.

Rate constants k_1, k_{-1}, and k_{cat} are used to define the MM model. The second-order associated forward rate constant for substrate binding k_1 describes ES association from the total substrate and free enzyme; k_{-1} is a second-order reverse rate constant for ES dissociation and k_{cat} is the first- order catalytic rate constant of the ES transition state and substrate turnover into the product (Srinivasan, 2022). K_D is a dissociation constant calculated using only binding and unbinding constants (k_{-1}/k_1). K_m^c can be calculated by $(k_{-1} + k_{cat})/k_1$ (Equation 2.6). Therefore K_m^c deviates from K_D due to k_{cat}. For simplicity, it was assumed that ES dissociation is more rapid than the turnover of the substrate into the product, and therefore K_m^c equated to K_D. The k_{cat} is a part of the lumped parameter v_{max}, representing the rate of product formation when the enzyme is saturated and reflects the enzyme's maximum rate.

Apparent kinetic parameters of reactions catalysed by CYP153A6 and CYP153A13 whole cells on *n*-octane, *n*-nonane, and *n*-decane were determined from the experimental data of two biological repeats, BR1, BR2 and BR3. Product concentration as a function of time data was directly fitted to the MM model (Equation 4.1) to estimate apparent kinetic parameters using least squares NLR (Gardossi et al., 2010, Duggleby, 1995) (Chapter 3.4.1.2). Least squares NLR minimised the combined residual sum of squared error of the difference between the MM model and the experimentally measured product concentration. Kinetic parameters v_{max} and K_m^c were estimated by minimising the objective sum of squared residuals (SSR) function ($SSR = \sum_i([P]_{t,exp} - [P]_{t,model})^2$ using Solver add-in on Microsoft Excel (Kemmer and Keller, 2010).

The K_D data determined by Funhoff et al. (2006) and Bordeaux et al. (2011) were respectively used as initial estimates to fit experimentally determined product concentration over time collected in the present study to the Michaelis-Menten model equation (Tables 4.11 and 4.12). It was noteworthy that these K_D estimates reported by authors Funhoff et al. (2006) and Bordeaux et al. (2011) were determined using purified enzymes, whereas the present study used whole cell biocatalysts. Therefore, these initial estimates introduced some limitations and required assumptions. It was assumed that the MM constants obtained by previous studies using purified enzymes could be used as initial estimates in the NLR procedure for whole cell biocatalyst systems. Therefore, the apparent MM constants of different alkane substrates (C_8, C_9, C_{10}) were compared for each whole cell biocatalyst and the affinity of the enzymes for these substrates were compared. It should be noted that the comparisons made were of the general trends between substrates by each enzyme, rather than statistically significant comparisons, due to the error associated with the NLR model predictions and experimental measurements.

Table 4.11 Initial estimates of Michaelis-Menten kinetic parameters of CYP153A6 whole cells on longer chain alkanes used for the fitting of data to the model using least squares nonlinear regression

Kinetic parameter	Alkane	BR1	BR2
v_{max} ($\mu mol_{1-alkanol} \cdot \mu mol_{P450} \cdot min^{-1}$)	*n*-octane	13.00	0.600
	n-nonane	2.827	0.238
	n-decane	n.d	0.481
K_m^c (μM)	*n*-octane	0.17	0.17
	n-nonane	0.022	0.022
	n-decane	n.d	0.01

n.d=not determined

Table 4.12 Initial estimates of Michaelis-Menten kinetic parameters of CYP153A13 whole cells on longer chain alkanes used for the fitting of data to the model using least squares nonlinear regression

Kinetic parameter	Alkane	BR1	BR2
v_{max} ($\mu mol_{1-alkanol} \cdot \mu mol_{P450} \cdot min^{-1}$)	*n*-octane	0.874	0.036
	n-nonane	0.407	0.028
	n-decane	n.d	0
K_m^c (μM)	*n*-octane	29	29
	n-nonane	n.d	n.d
	n-decane	n.d	16

Equation 3.6 ($v_{max} = k_{cat}[E]$) (Chapter 3.4.1.2) was used to obtain the initial NLR v_{max} estimates by substituting k_{cat} with the maximum TOF. Total enzyme concentration was drawn from the CO-difference spectra measurement before reaction initiation, and it was assumed that enzyme deactivation did not occur throughout the 72-hour reaction. Previous biocatalytic studies only reported rates in terms of activity or TOF of CYP153 rather than the v_{max}. The latter strictly refers to saturation kinetics and reaction rates under complete ES saturation, in which the reaction rate is in zero-order and does not depend on the substrate concentration (Kozuch and Martin, 2012). Therefore, since there was a lack of reported v_{max} rates in previous literature due widespread reporting of TOF rather than the Michaelis-Menten kinetic parameters, NLR v_{max} results were compared to the TOFs reported by previous authors at single-time points (Tables 4.13 and 4.14). Furthermore, previous CYP153 characterisation studies determined the K_D constants by measuring a spectroscopic shift of the CYP153's active site ferric heme iron from a low-spin to a high-spin state upon alkane substrate binding (Funhoff et al., 2006, Bordeaux et al., 2011, Ferrario et al., 2018).

The v_{max} of CYP153A6 on *n*-octane was 3-fold and 16-fold faster in BR1 (0.008 $\mu mol \cdot L_{BRM}^{-1} \cdot min^{-1}$) than in BR3 and BR2, respectively (Table 4.13). The v_{max} of CYP153A13 on *n*-octane the fastest (0.012 $\mu mol \cdot L_{BRM}^{-1} \cdot min^{-1}$) in BR3, followed by BR1 and then BR2 (Table 4.14). The v_{max} achieved by CYP153A6 and CYP153A13 on *n*-nonane was 6 and 10-fold faster in BR1 than in BR2, respectively. The v_{max} of CYP153A6 on *n*-nonane was 3-fold faster in BR1 than in BR3. No product was produced by CYP153A6-decane, CYP153A13-nonane and CYP153A13-decane in BR3 because of delayed onset of reaction (discussed earlier in Chapter 4.2) and therefore their maximum reaction velocities were not determined. The variability obtained between biological repeats was discussed in Chapter 4.2. Furthermore, the estimated v_{max} results were vastly different because in BR1, data points from 0 to 72 hours were included, BR2 only included data points from 24 to 72 hours and BR3 included data collected from 0 to 12 hours.

Moreover, the difference between the v_{max} rates of *n*-octane and *n*-nonane was higher in BR1 than in BR2. As stated previously, least squares NLR for BR1 utilised product concentration data from 3 to 72 hours, whereas BR2 only included data from 24 to 72 hours. BR1 included the initial phases of the reaction, in which delays in reaction initiation with substrates of lower aqueous solubility were observed. Therefore, the inclusion of this initial reaction phase data for least squares NLR modelling in BR1, resulted in a greater difference between the substrate's v_{max} rates in BR1 than in BR2. Initial reaction data in BR3 also showed a greater difference between the v_{max} rates on *n*-octane and *n*-nonane because it only included data from 0 to 12 hours of the reaction, due to late onset of reaction seen with the less soluble *n*-nonane.

Table 4.13. Apparent Michaelis-Menten kinetic parameters (v_{max}, K_m^c) of CYP153A6 whole cell biohydroxylation of medium-chain alkanes determined using least squares nonlinear regression compared to previous studies' TOF data at single-time points and binding constants.

Parameter	Biological repeat or study	*n*-octane	*n*-nonane	*n*-decane
Specific activity or TOF ($\mu mol_{1-alkanol} \cdot \mu mol_{P450}^{-1} \cdot min^{-1}$) [a]	BR1[d]	16.1	5.1	n.d
	BR2[d]	1.28	0.28	0.40
	BR3[d]	3.95	0.83	0
	Funhoff et al. (2006) [e]	60.8	49.7	13.8
	Funhoff et al. (2007) [f]	58	n.a	26
	Pennec et al. (2015) [g]	9.2	n.a	0.2
v_{max} ($\mu mol \cdot L_{BRM}^{-1} \cdot min^{-1}$) [b]	BR1[d]	0.008	0.002	n.d
	BR2[d]	0.0005	0.0003	0.0004
	BR3[d]	0.003	0.0006	n.d
K_m^c (μM)[b]	BR1[d]	0.17	0.022	n.d
	BR2[d]	0.17	0.022	0.010
	BR3[d]	0.17	0.022	n.d
K_D[c]	Funhoff et al. (2006) [e]	0.17	0.022	0.010

[a]: Activity was reported in previous literature at single-time points at 24 hours and reaction conditions (see [e-g] for conditions). In the present study (see [c] for conditions), the ***maximum*** TOF was determined in Chapter 4.2 measured in $\mu mol_{1-alkanol} \cdot \mu mol_{P450} \cdot min^{-1}$

[b]: The current study's maximum reaction velocity (v_{max}) and the MM constant (K_m^c) calculated fitting experimental product concentration as a function of time data to the MM model using NLR

[c]: The present study's conditions- *E. coli* whole cell biocatalyst; 2LPS; 20˚C

[d]: Purified CYP153A6; Expressed in natural host; product concentration measured at 0.75 hours; 2LPS; 30˚C

[e]: Purified CYP153; Expressed in natural host; product concentration measured at 0.75 hours 2LPS; 30˚C

[f]: *E. coli* whole cell biocatalyst; product concentration measured at 8 hours; 20˚C

[g]: n.d is not determined because no product concentration was detected within the specified reaction time

[h]: n.a means not applicable because author did not perform tests on substrate

Table 4.14 Apparent Michaelis-Menten kinetic parameters (v_{max}, K_m^c) of CYP153A13 whole cell biohydroxylation of medium-chain alkanes determined using least squares nonlinear regression compared to previous studies' turnover frequency data at single-time points and binding constants.

Parameter	Biological repeat or study	*n*-octane	*n*-nonane	*n*-decane
Specific activity or TOF$\mu mol_{1-alkanol} \cdot \mu mol_{P450}^{-1} \cdot min^{-1}$ [a]	BR1[d]	3.8	1.8	n.d
	BR2[d]	0.21	0	0
	BR3[d]	1.56	0	0
	Funhoff et al. (2007) [f]	0	n.a	6
	Pennec et al. (2015) [g]	1.7	n.a	0.2
	Bordeaux et al. (2011) [h]	57	n.a	30
v_{max} ($\mu mol_{1-alkanol} \cdot L_{BRM}^{-1} \cdot min^{-1}$) [b]	BR1[d]	0.0008	0.0003	n.d
	BR2[d]	0.00004	0.00003	0
	BR3[d]	0.012	n.d[i]	n.d[i]
K_m^c (μM)[b]	BR1[d]	29	n.d[i]	16
	BR2[d]	29	n.d[i]	16
	BR3[d]	29	n.d[i]	n.d[i]
K_D[c]	Bordeaux et al. (2011) [h]	29	n.a	16

[a-f]: See notes below Table 4.13

[g]: Purified CYP153A13; Monophasic; 25˚C; product concentration measured at 0.17-hr reaction time

[h]: n.d is not determined because no product concentration was detected within the specified reaction time

[i]: n.a means not applicable because author did not perform tests on substrate

Due to the variability of maximum reaction velocities between biological repeats, comparisons were made within each biological repeat. According to the least squares NLR modelling results, the higher solubility of *n*-octane compared to *n*-nonane facilitated faster maximum reaction velocities by whole cell biocatalysts, respectively (Tables 4.13 and 4.14). In BR1, the fastest v_{max} achieved by CYP153A6 whole cells was on *n*-octane with a rate of 0.008 $\mu mol_{1-octanol} \cdot L_{BRM}^{-1} \cdot min^{-1}$, which was four-fold faster compared to *n*-nonane (0.002 $\mu mol_{1-nonanol} \cdot L_{BRM}^{-1} \cdot min^{-1}$) (Table 4.13). In BR2, these results were validated with the fastest v_{max} of 0.0005 $\mu mol_{1-octanol} \cdot L_{BRM}^{-1} \cdot min^{-1}$ achieved by CYP153A6 on *n*-octane, which was approximately two-fold faster compared to *n*-nonane with a v_{max} of 0.0003 $\mu mol_{1-nonanol} \cdot L_{BRM}^{-1} \cdot min^{-1}$ (Table 4.13). In BR3, a similar trend was observed with a v_{max} of 0.003 $\mu mol_{1-octanol} \cdot L_{BRM}^{-1} \cdot min^{-1}$, which was five-fold faster than on *n*-nonane (0.003 $\mu mol_{1-octanol} \cdot L_{BRM}^{-1} \cdot min^{-1}$).

The same trend was observed with reactions catalysed by CYP153A13 (Table 4.14). In BR1, the v_{max} of 0.0008 $\mu mol_{1-octanol} \cdot L_{BRM}^{-1} \cdot min^{-1}$ achieved by CYP153A13 whole cells on *n*-octane was nearly three-fold faster than *n*-nonane with a v_{max} of 0.0003 $\mu mol_{1-nonanol} \cdot L_{BRM}^{-1} \cdot min^{-1}$. Similarly, in BR2,

the v_{max} rate achieved by CYP153A13 whole cells was 1.3-fold faster on *n*-octane with a v_{max} of 0.00004 $\mu mol_{1-octanol} \cdot L_{BRM}^{-1} \cdot min^{-1}$ than *n*-nonane with a v_{max} of 0.00003 $\mu mol_{1-nonanol} \cdot L_{BRM}^{-1} \cdot min^{-1}$ (Table 4.14).

These results were explained by the same reasons provided to explain the relative magnitude of maximum TOFs obtained for *n*-octane compared to *n*-nonane in Chapter 4.2. Transport across the cellular membrane and LL interface becomes more limiting as the aqueous solubility of the substrate decreases. Furthermore, in BR2, CYP153A6 whole cells achieved a v_{max} two-fold higher on *n*-octane than on *n*-decane (Table 4.13). The correlated relationship between v_{max} and solubility was established from the results in Tables 4.13 and 4.14, but there was an exception. CYP153A6 whole cells unexpectedly achieved a 1.3-fold faster maximum velocity rate on *n*-decane (0.0004 $\mu mol_{1-decanol} \cdot L_{BRM}^{-1} \cdot min^{-1}$) than on *n*-nonane (0.0003 $\mu mol_{1-nonanol} \cdot L_{BRM}^{-1} \cdot min^{-1}$), even though *n*-decane is less soluble in the aqueous phase than *n*-nonane. As discussed in Chapter 4.2, these results reflected a higher reaction rate on *n*-decane than *n*-nonane due to the odd-even effect of liquid alkanes. As mentioned previously, the odd-even effect of liquid alkanes could have resulted in difficulties in transporting odd-chain alkanes across the cellular membrane or LL interface. Transmembrane transport was potentially impacted by the odd-even of alkanes effect from entrapment of *n*-nonane in the cellular membrane due to its anisotropic properties that contribute to less ordered packing efficiency. Inaccuracy of the v_{max} NLR results may have resulted from the assumption that total enzyme concentration remains consistent over time because whole cells were exposed to toxic solvents, substrates, and products, as shown in Chapter 4.2 and a study by Olaofe et al. (2013). This assumption may result in the overestimation of NLR v_{max} results.

Regardless of the initial estimates selected, the v_{max} NLR results repeatedly resolved to the same results. The K_m^c constants obtained from least squares NLR did not show deviation from the respective initial estimates drawn from previous literature (Tables 4.11 and 4.12). This finding revealed that the least squares NLR procedure to estimate K_m^c values relied heavily on the accuracy of values drawn from previous studies as initial estimates. Altogether, these findings supported previous findings by Lauchnor et al. (2015), who analysed the sensitivity of the optimised NLR kinetic parameters and found that fitting the MM model to experimental whole cell biocatalytic data was more sensitive to changes in v_{max} rates than K_m^c constants. Overall, uncertainty existed in the estimation of K_m values for whole cell biocatalysis, which was in agreement with Lauchnor et al. (2015).

Due to its complexity, few studies have developed an approach to determining the kinetics of whole cell biocatalysis reactions in a 2LPS. Whole cell K_m^c constants depict a lumped parameter that accounts for ES affinity, substrate-solvent interactions, and transmembrane transport (Lauchnor et al., 2015, Ferrario et al., 2018). Potential inaccuracies in this approach include the assumption that K_D constants obtained for purified forms of CYP153A6 and CYP153A13 can be used for the NLR procedure to determine K_m^c values for whole cells in a 2LPS. The NLR procedure likely overestimated K_m^c constants because substrate affinities for whole cells are lower than for isolated enzymes (Lauchnor et al., 2015).

The determination of K_m^c constants of alkane substrates of poor aqueous solubility in CYP153 whole cell 2LPSs was challenging, confirming previous studies' findings (Nie et al., 2014b, Funhoff et al., 2006). These results confirmed previous sentiments by Cooley et al. (2009), who stated that the slow substrate diffusion rates and limited detection of submicromolar concentrations by a GC presented a challenge in determining the K_m of CYP enzymes on medium to longer chain alkanes.

Although there was uncertainty in the accuracy of the K_m^c constants, these constants decreased as alkane chain length increased and solubility decreased (Tables 4.13 and 4.14). Since NLR K_m^c constants

were almost identical to the selected initial estimates, the results reflected those of previous studies, showing that longer-chain alkanes bind more strongly to the active site of the CYP153 enzyme (Lee et al., 2016, Funhoff et al., 2006, Bordeaux et al., 2011). Initial estimates were selected from previous studies by Funhoff et al. (2006) and Bordeaux et al. (2011) who determined the dissociation constants K_D of CYP153A6 and CYP153A13 towards various linear alkanes by determining the shift in UV-visible spectra of samples upon enzyme binding to the alkane. In the field of biocatalysis, the MM constants for the hydroxylation of linear alkanes by CYP153 whole cell biocatalysts have received very little attention. Therefore, initial estimates had to be based on dissociation constants from the studies by Funhoff et al. (2006) and Bordeaux et al. (2011), even though these values were determined for purified enzymes.

Longer alkane or alcohol chain lengths possess higher hydrophobicity, resulting in tighter binding to an active site, confirming previous evidence by Ji et al. (2013) and van Beilen et al. (2003). These findings were attributed to the strong molecular interactions between the hydrophobic alkane substrate and the hydrophobic pocket of the active site of the CYP153 enzymes, also known as the 'hydrophobic effect' (Funhoff et al., 2006). Slow or ineffective release of the hydrophobic alcohol also decreases the frequency of enzyme-substrate binding which becomes rate-limiting (Cooley et al., 2009).

These results and previous studies highlight an interesting phenomenon of longer chain alkane or alcohol product binding very tightly to the active site, to the extent that it limits the reaction rate, as illustrated by the decreased v_{max} rates achieved by CYP153 whole cell biocatalysts (Tables 4.13 and 4.14). Therefore, decreases in v_{max} rates as chain length increases related to the tight substrate or product binding that becomes rate-limiting due to the hydrophobic effect. These findings were supported by Park et al. (2022), product release is more rate-limiting than substrate turnover into the product.

These findings corroborated with results by Bordeaux et al. (2011), who also illustrated that the degree of electron coupling decreases as hydrophobicity or aqueous solubility of the alkane increases due to an increase in its chain length. Furthermore, a decrease in electron coupling could result from the inadequate or slow release of the hydrophobic substrate or product from the active sites, both of which would decrease the availability of the active site to bind new substrate in the proximal area, as shown by van Beilen et al. (2003). This results in lower electron coupling which causes an increase in oxidative side reactions to produce undesired products such as H_2O_2 and consequently results in slower v_{max} rates (Bordeaux et al., 2011, van Beilen et al., 2003a). To determine how well the data fit the model, R^2 coefficients were calculated Equation 3.12 in Chapter 3.4.1.2 ($R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}$)(Nagar et al., 2014). High R^2 values (≥0.9) were obtained for the NLR results which meant that the data fitted the MM model relatively well (Table 4.15).

Table 4.15. R squared values to indicate how well the experimental data fits the Michaelis-Menten model

	BR1	BR2
CYP153A6-octane	0.92	0.96
CYP153A13-octane	0.94	0.98
CYP153A6-decane	n/a	0.97
CYP153A6-nonane	0.93	0.96
CYP153A13-nonane	0.90	0.84

The results repeatedly suggested that substrate transport across the cellular membrane and LL interface was limited due to low substrate solubility. Substrate concentration available in the aqueous phase was reliant on the thermodynamic phase equilibrium of the 2LPS, justifying the need to apply excess substrate to maximise substrate interfacial transfer and the use of an organic carrier solvent that allows further substrate solubilisation in the aqueous phase (Olaofe et al., 2013, Kim et al., 2007). Factors such as substrate transport across the LL interface and membrane create a low cytoplasmic substrate concentration in the proximity of the active site such that $[S] \ll K_m$. Therefore, the assumption of the MM model concerning substrate saturation was likely invalid and that zero-order kinetics were not reached because $[S] \ll K_m$ (Kozuch and Martin, 2012, Nagar et al., 2014). Accordingly, it is argued in Chapter 4.4 that CYP153A13 whole cells did not show maximum functionality or saturation kinetics in the whole cell 2LPSs under study as it was limited by substrate concentration, while CYP153A6 may have shown saturation kinetics and was not limited by substrate concentration.

In summary, apparent kinetic parameters facilitated an understanding of factors that may be rate-limiting in the whole cell 2LPS, assuming that zero-order saturation kinetics is reached in relation to the substrate. These factors ranged from enzyme expression levels to the physicochemical properties (such as substrate solubility) of the 2LPS. The findings of the least squares NLR were similar to the TOF over time analyses in Chapter 4.2. For example, the expression strategy of CYP153A13 requires improvement to achieve faster v_{max} rates.

There was substantial evidence that an increase in substrate solubility allows an increase in v_{max} rates, similarly, shown with the TOF maxima in Chapter 4.2. Unexpectedly, higher v_{max} rates were achieved by CYP153A6 whole cells on *n*-decane than *n*-nonane, potentially because of the odd-even effect of liquid alkanes. CYP153A13 whole cells showed no activity on *n*-decane, which illustrated that insufficient enzyme expression negatively impacted the v_{max} to a greater extent than substrate partitioning into the aqueous phase or transmembrane transport, since CYP153A6 showed activity on *n*-decane.

The role of the hydrophobic effect on CYP153's reaction mechanism on medium-chain alkanes was highlighted. As substrate or product solubility decreases, the binding affinity increases between the active site and hydrophobic molecule, preventing active site availability and limiting the reaction. The low aqueous solubility of longer-chain alkanes presented difficulties in determining their kinetic parameters, confirming the findings of several studies. The determination of kinetic parameters was further complicated using whole cells, which likely resulted in the overestimation of v_{max} rates and K_m^c constants. The least squares NLR modelling approach showed more sensitivity to changes in the v_{max} compared to K_m^c. Besides these challenges, the least squares NLR approach itself presented its difficulties, yet high R^2 values were obtained, indicative of a good fit of the data to the MM model.

Alternative initial estimates were tested in an attempt to improve model sensitivity, but these did not result in different predicted kinetic parameters.

Notably, apparent kinetic parameters include molecular interactions such as hydrogen bonds between the substrate and the active site of enzymes and those between the substrate and its surrounding molecules, i.e., solvent (Grosch et al., 2017, Ferrario et al., 2018). Therefore, a low K_m^c constant indicates either favourable affinity between the enzyme and substrate or unfavourable interactions between substrate and solvent. The term 'apparent' refers to the assumption that enzymes were under total exposure to substrate concentration. Apparent kinetics did not account for the nature of the 2LPS, in which concentrations change over time as volumes changes and substrate availability depends on partitioning and thermodynamic activity. Therefore, using the thermodynamic activity of components is advantageous instead of concentration. Further investigation was performed to depict the true kinetics of CYP153 whole cell biohydroxylation catalysis in a 2LPS using thermodynamic activities and accounting for solvation effects.

4.4 Solvation-corrected kinetics of 2LPS CYP153 whole cell biocatalysis of primary hydroxylation of C_8, C_9 and C_{10} alkanes

In Chapter 4.3, the experimental kinetics of whole cell biohydroxylation of medium chain alkanes were investigated using experimentally measured product concentration as a function of time and nonlinear regression analyses to determine the apparent Michaelis-Menten (MM) constant, K_m^c, for each enzyme-substrate (ES) pair. The K_m^c constant reflects the molecular interactions between ES complex and the substrate-solvent interactions within non-ideal mixtures. Substrate-solvent interactions introduce deviations in the accurate assessment of true ES affinity (Ferrario et al., 2018).

To overcome shortcomings of the above approach, in Chapter 4.4 the non-ideal nature of the system and the deviations introduced by substrate-solvent interactions are addressed using an approach developed by Ferrario et al. (2018) (*viz*. Chapter 3.5.2). Solvation effects were investigated by studying the impact of substrate-solvent interactions in the presence of alkanes with increasing chain length which, in turn, affect the solubility (Pleiss, 2017). Thermodynamic activity-based MM constants (K_m^a) were determined; these are independent of SS interactions or solvation effects and thereby reflect the true affinity between an enzyme and substrate by excluding the effects from the molecular interactions arising between substrate and solvent as well as transport effects impacting the whole cell catalyst.

The K_m^a constants of each ES pair were determined to reveal accurate shape complementarity and affinity between the respective substrate and enzymes to clarify the linear alkane substrate scope of CYP153A6 and CYP153A13. To develop large scale biocatalytic systems, it is necessary to understand 'one-pot' systems that comprise multiple enzymes or substrates. In one-pot systems, low substrate supply favours the catalysis of the reaction utilising the ES pair with the lowest K_m constant, highlighting the importance thermodynamic activity-based K_m^a constants (Kim et al., 2007).

4.4.1 The determination of thermodynamic activity coefficients of alkanes

The thermodynamic activity coefficient (γ_i^∞) of each alkane was determined and used to convert K_m^c constants (determined in Chapter 4.3) to thermodynamic activity-based constants (K_m^a). Given that 2LPSs are non-ideal solutions, the activity coefficients of *n*-octane, *n*-nonane, and *n*-decane were respectively determined in the aqueous phase to ascertain the deviation of the alkane from its ideal state (Janssen et al., 1995). Experimental determination of activity coefficients is costly and time-consuming, therefore predictive, and theoretical methods were employed to simulate the experimental 2LPS. These methods enabled the determination of activity coefficients of alkanes at thermodynamic equilibrium (Gerber and Soares, 2010).

For the predictive methods, a process flowsheet was developed on Aspen Plus (V12) to simulate the equilibrium state of wet-laboratory 2LPS conditions utilised for whole cell biocatalytic reactions at 20˚C. This enabled the determination of equilibrium-related parameters which could be used for further analyses to gain understanding of biocatalytic 2LPSs. The process flowsheet comprised three-unit operations and component streams specified to simulate the 2LPS under study.

Two input streams (ALKANE and BEHP) were connected to a mixer (MIXER1), allowing mixing of the organic phase stream. One stream consisted of *n*-alkane (ALKANE) and flowed into the mixer at a flow rate of 0.0002 $L \cdot min^{-1}$ and the other consisted of BEHP organic solvent carrier (BEHP), and flowrate was 0.0001 $L \cdot min^{-1}$. These flow rates were selected to mimic the aqueous-to-organic phase ratio of 1:0.3 and the substrate-to-organic solvent ratio of 2:1 used in the experimental studies. The ALKANE and BEHP streams combined in MIXER1 to form the organic liquid phase, an outlet stream named ORG1.

ORG1 outlet stream and a stream consisting of only distilled water (AQ) at a flow rate of 0.001 $L \cdot min^{-1}$ were connected to a second mixer (MIXER2) to form a two-liquid phase stream. The two-liquid phases

in MIXER2 were passed through an outlet stream (LLE) and sent into a single-stage separator (DECANTER) to separate the two liquid phases based on the UNIF-LL model. The UNIF-LL model is commonly used to predict thermodynamic behaviour of biocatalytic 2LPSs, such as alkane partitioning behaviour between two liquid phases. UNIF-LL models also allowed the prediction of activity coefficients in systems that lack experimentally determined phase equilibrium data (Fredenslund et al., 1975).The DECANTER split the two-phase liquids by minimising the Gibbs free energy of the system. The resultant streams LIQ1 and LIQ2 represented the 2LPS at equilibrium. The liquid phase that resulted in the higher mole fraction of pure water was denoted as the aqueous phase, and the liquid phase that consisted of primarily organic co-solvent (BEHP) and alkane was denoted as the organic phase (Chapter 3.5.1). In addition to UNIF-LL process flowsheet models simulating experimental conditions at 20˚C, UNIF-LL simulations were performed at 25˚C to allow comparison to the activity coefficients reported in previous literature (Ferrario et al., 2018).

For the theoretical method, the activity coefficients could be determined at infinite dilution, due to the poor water solubility of the alkanes of interest, using Equation 3.15 ($c^{sol} = 55.5\ M/\gamma^{\infty}$) (Halling, 1994, Ferrario et al., 2018, Banerjee, 1985). Therefore, the solubility limit (c^{sol}) of each alkane and the mole-based concentration of pure liquid water (55.5 M) could be used to determine the activity coefficients at infinite dilution (Chapter 3.5.1.1).

The accuracy of the predicted activity coefficients obtained from the process flowsheet using the UNIF-LL model were validated. A parity plot was constructed to compare predicted activity coefficients with previously determined experimental values by Abraham et al. (2017) of *n*-octane, *n*-nonane and *n*-decane (Figure 4.11). In assessing the validity of a model, experimental data (observed) determined by Abraham et al. (2017) were plotted against predicted data (computationally simulated). The relationship between observed and simulated should ideally show a straight line that goes through the plot origin to provide a correlation coefficient (R^2) value of 1. The higher the data parity between experimentally determined and predicted or theoretically calculated activity coefficients (respectively), the higher tendency of the correlation coefficient (indicated by an R^2 value) towards one and the closer the data to the parity line. As a reminder, predicted or simulated data shown in Figure 4.11 was determined using the UNIF-LL model developed in Aspen Plus (V12), while theoretical data refers to the activity coefficient determined using Equation 3.15.

Activity coefficients, assuming infinite dilution, that were theoretically determined from Equation 3.15 were also compared with experimental values by Abraham et al. (2017) and predicted values plotted in Figure 4.11. To increase the legibility and clarity of the parity plot, the magnitude of the activity coefficient values was reduced by applying the logarithm function to each individual activity coefficient value of the data set (Figure 4.11).

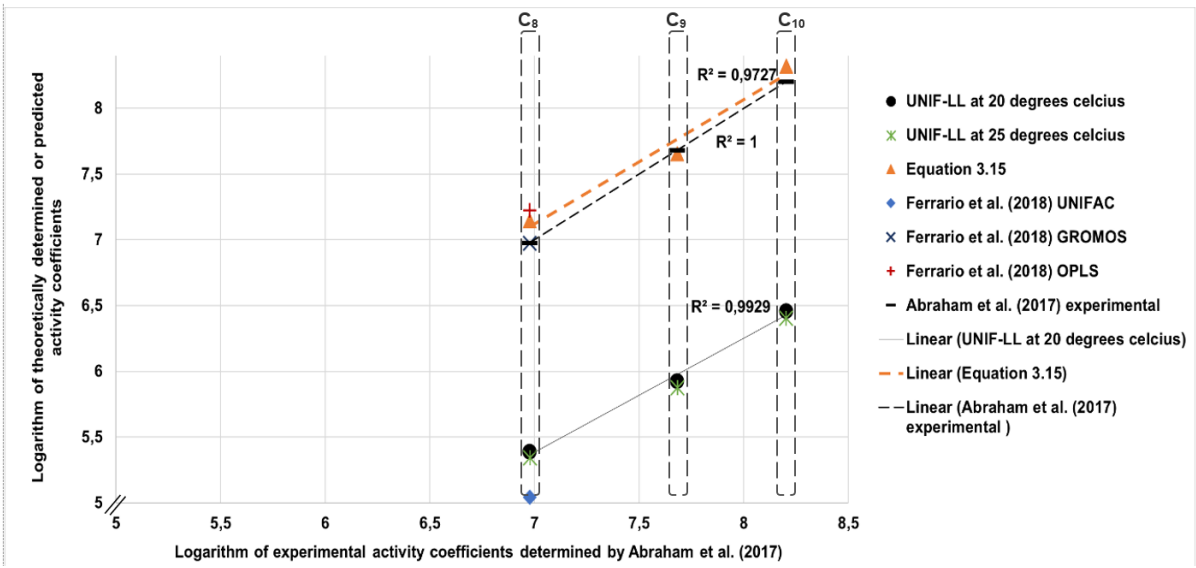


Figure 4.11 A parity plot illustrating the relationship between experimentally determined log γ^{∞}values (by Abraham et al. (2017) and predicted and theoretically determined log γ^{∞}values of n-octane (far left), n-nonane (middle), and n-decane (far right), respectively. Each data point plotted represents an (x; y) coordinate, with x as the experimentally determined values and y as the predicted or theoretically calculated values. For experimental data determined by Abraham et al. (2017) (indicated by ■), x=y to obtain an R2 of 1 and a dashed (---) parity line. The correlation between experimentally determined activity coefficients (x-axis) determined by Abraham et al. (2017) and predicted (UNIF-LL) or theoretically calculated (Equation 3.15) or previously determined activity coefficients (y-axis) was illustrated. All activity coefficients, unless stated otherwise, were determined at 25˚C in the aqueous phase.

Activity coefficients determined using Equation 3.15 (indicated by ▲) showed a high correlation with values experimentally determined by Abraham et al. (2017), with an R^2 value of 0.9727 as shown in Figure 4.11. Therefore, the activity coefficients of *n*-octane, *n*-nonane and *n*-decane determined using Equation 3.15 were compatible in magnitude to the activity coefficients experimentally determined by Abraham et al. (2017), with $log\,\gamma^{\infty}$ values ranging from approximately 7 (for *n*-octane) to 8.3 (for *n*-decane). The high correlation makes sense because Abraham et al. (2017) determined activity coefficients of alkanes under the same conditions (infinite dilution in pure water) when using Equation 3.15.

Similarly, the experimental activity coefficients determined by Abraham et al. (2017) were consistent with activity coefficients for *n*-octane predicted by molecular dynamic simulations (GROMOS and OPLS indicated by + and ×, respectively) carried out by Ferrario et al. (2018) (Figure 4.11). This high correlation is explained by the fact that both authors determined activity coefficients at infinite dilution in pure water.

At 20˚C and 25˚C the activity coefficients predicted by UNIF-LL (indicated by ● and ✕) were of comparable magnitude to each other (activity coefficients were only 0.82% higher at 20˚C compared to 25˚C), which means that a change in temperature did not impact the activity coefficient of the alkane substantially (Figure 4.11). The UNIF-LL process flowsheet model was validated because they were comparable to the UNIFAC coefficients predicted by Ferrario et al. (2018) (Figure 4.11, ♦). Figure 4.11 illustrates that in comparison to experimentally determined $log\,\gamma^{\infty}$ values by Abraham et al. (2017), UNIF-LL predicted $log\,\gamma^{\infty}$ values of each alkane (ranging from 5.5 for *n*-octane to 6.5 for *n*-decane) were underestimated by 22%. These findings were substantiated by Ferrario et al. (2018) who showed that GROMOS MD simulations were in the correct order of magnitude to experimentally determined

activity coefficients by Abraham et al. (2017), while activity coefficients predicted by the UNIFAC model were underestimated. Ferrario et al. (2018) stated that UNIFAC predictions underestimated activity coefficients compared to MD simulations (Figure 4.11).

In other words, the deviation of the alkane's activity coefficient from ideality was underestimated by UNIF-LL models. The UNIF-LL process flowsheet model predicted lower activity coefficients and weaker molecular interactions than reality because it does not fully account for molecular interactions between the alkane and aqueous solvent. Previous studies found that UNIFAC-derived activity coefficients at infinite dilution resulted in underestimated water solubility of compounds (Banerjee, 1985, Arbuckle, 1983). This indicates the alkane-aqueous molecular interactions were stronger than UNIF-LL had predicted.

UNIF-LL models tend to underestimate activity coefficients for larger hydrophobic substrates (Banerjee, 1985). This is due to the fact that larger compounds possess a high number of chemical functional groups, resulting in an increase in overall accumulated error in the additivity from the slight deviations of each chemical group (Arbuckle, 1983, Ferrario et al., 2018, Banerjee, 1985). The outcomes confirmed that UNIF-LL models underestimate the activity coefficients in aqueous solution for hydrophobic compounds with larger molecular sizes, such as C_8, C_9, and C_{10} alkanes.

In agreement with previous studies, Figure 4.11 demonstrated that the activity coefficients derived from the UNIF-LL model increased as the solubility of alkanes decreased (Abraham et al., 2017, Ferrario et al., 2018). This makes sense in light of thermodynamic equilibrium, in which a compound's thermodynamic activity is identical in each phase of the 2LPS. This demonstrates the relationship between activity coefficient and concentration whereby lower activity coefficients are generally associated with a higher concentration (or solubility in this case) and vice versa (Halling, 1994).

The process flowsheet simulation results could be improved by using modified versions of the UNIF-LL (Ferrario et al., 2018). In comparison to UNIF-LL models, COSMO-RS models, which are based on computational quantum mechanics, have been shown to provide highly accurate estimates of activity coefficients for biocatalytic 2LPSs and are recommended for future study (Grosch et al., 2017, Navas et al., 2010, Klamt, 2005, Gerber and Soares, 2010). Nevertheless, with the caveat that the UNIF-LL model tends to underestimate activity coefficients by 22% for *n*-octane, *n*-nonane and *n*-decane, respectively, and that the simulation only took into account the initial state of the 2LPS prior to the introduction of the whole cell biocatalyst, the results of the UNIF-LL may still serve as an estimate for the substrate partitioning behaviour and substrate-solvent interactions within the 2LPS. Therefore, the activity coefficients derived from the UNIF-LL model and the additional theoretical approach using Equation 3.15 were used to determine the thermodynamic based Michaelis-Menten constants.

4.4.2 Thermodynamic activity-based Michaelis-Menten constants (K_m^a)

Aqueous phase activity coefficients can be used to determine thermodynamic activity-based constants K_m^a, which reflect the true affinity between an enzyme and substrate. In Chapter 4.3, K_m^c constants were determined in nonideal conditions, which reflects both enzyme-substrate and substrate-solvent molecular interactions. Interactions between substrate and aqueous molecules can impact the substrate's ability to bind to CYP's active site and therefore affects the effective concentration of the substrate available to the active site. Therefore, the determination of the true affinity between enzyme and substrate is affected by interference arising molecular interactions between the substrate and the solvent. Substrate-solvent interactions are excluded by calculating the K_m^a constant, thereby revealing the true affinity between substrate and enzyme.

K_m^a constants were calculated using Equation 3.16 ($K_m^a = \gamma^\infty \cdot K_m^c/55.5\ M$) (Chapter 3.5.2). The activity coefficients (displayed in Figure 4.11) and K_m^c constants (*the* latter determined in Chapter 4.3) were substituted into Equation 3.16 for each ES pair (Tables 4.16 and 4.17). The K_m^c constants of CYP153A6 and CYP153A13 were estimated by least squares nonlinear regression analyses by fitting product concentration data as a function time data to the Michaelis-Menten model in Chapter 4.3. Initial estimates for K_m^c used in the nonlinear regression analyses were taken from Funhoff et al. (2006) and Bordeaux et al. (2011) (Tables 4.11 and 4.12).

The activity coefficients shown in Figure 4.11 were determined using two different methods: (1) theoretically using Equation 3.15 and (2) computationally using the UNIF-LL process flowsheet model. These activity coefficients are both used to determine thermodynamic activity-based MM constants. The study focused the relevance of considering thermodynamic activity when investigating the kinetic analysis of whole cell biocatalyst alkane activation; this has received little attention in prior research.

By substituting activity coefficients (determined in Chapter 4.4.1) and K_m^c constants (determined in Chapter 4.3) into Equation 3.16, the K_m^a constants of each enzyme pair were determined in Chapter 3.5.2). In Tables 4.16 and 4.17, the K_m^a constants are presented for each prediction and theoretical method on each substrate and enzyme. These are correlated to the K_m^c constants and alkane solubility. The difference between UNIF-LL and Equation 3.15 based K_m^a constants relate to previous findings that the UNIF-LL model underestimates activity coefficients by 22% compared to Equation 3.15's activity coefficients (*viz.* Figure 4.11 in Chapter 4.4.1).

Table 4.16. Activity-based MM constants determined using activity coefficients derived from UNIF-LL predictions or theoretical (using Equation 3.15) for CYP153A6 on n-octane, n-nonane, and n-decane, compared to K_m^c constants determined in Chapter 4.3 and solubility determined by Riddick et al. (1986).

	UNIF-LL based $K_m^a\ (\mu M)$	Equation 3.15 based $K_m^a\ (\mu M)$	$K_m^c(\mu M)$	Solubility (μM)
n-octane	0.000748	0.04288	0.17	4.00
n-nonane	0.000333	0.01783	0.022	1.22
n-decane	0.000515	0.03783	0.01	0.27

Table 4.17 Activity-based MM constants determined using activity coefficients derived from UNIF-LL predictions or theoretical (using Equation 3.15) for CYP153A13 on n-octane, n-nonane, and n-decane, compared to K_m^c constants determined in Chapter 4.3 and solubility determined by Riddick et al. (1986).

Substrate	$K_m^a\ (\mu M)$ from UNIF-LL derived γ^∞	Equation 3.15 based $K_m^a\ (\mu M)$	$K_m^c(\mu M)$	Solubility (μM)
n-octane	0.1276	7.3153	29	4.00
n-nonane	n.d	n.d	n.d	1.22
n-decane	0.8238	59.26	16	0.27

n.d= not determined because previous literature by Bordeaux et al. (2011) did not include CYP153A13-nonane and therefore no initial estimates were available for the determination of K_m^c constants by nonlinear regression analyses

The K_m^c constants of CYP153A6 and CYP153A13 towards *n*-octane were 227 times higher than the UNIF-LL based K_m^a constants, respectively, and only four-fold larger than Equation 3.15-based K_m^a constants (Tables 4.16 and 4.17). The K_m^c constants of CYP153A6 towards *n*-nonane was 67 and 1.2 times higher than the UNIF-LL and Equation 3.15-based K_m^a constants, respectively. UNIF-LL based K_m^a constants of CYP153A6 and CYP153A13 towards *n*-decane were 19-fold lower than their respective K_m^c constants, respectively (Tables 4.16 and 4.17).

According to the results so far, the K_m^c constants have underestimated the true enzyme-substrate affinity. This is clear from the K_m^a constants' lower values compared to the K_m^c constants. A deeper understanding of the true enzyme-substrate affinity was attained by excluding substrate-solvent interactions and considering only the interactions between the enzyme and substrate. This was accomplished by calculating K_m^a constants based on thermodynamic activity coefficients that solely account for enzyme-substrate interactions. Therefore, the results in Table 4.16 and 4.17 revealed that the enzyme-substrate affinity was greater than what was predicted by the K_m^c constants, which account for both substrate-solvent and enzyme-substrate interactions. This demonstrated the significant impact of solvent effects on the apparent affinity between an enzyme and substrate and also highlights the importance of accurately assessing the role of substrate-solvent interactions in biocatalytic reactions.

There was one exception to the results discussed in the above paragraph, in which Equation 3.15-based K_m^a constants of CYP153A6 and CYP153A13 towards *n*-decane were 3.8-fold higher than the K_m^c constants, respectively (Tables 4.16 and 4.17). According to these results, the true affinity between enzyme and substrate was overestimated by the K_m^c constant. This overestimation may be the result of the higher hydrophobicity of *n*-decane due to its longer chain length compared to *n*-octane and *n*-nonane, causing unfavourable molecular interactions with aqueous molecules. As a result, n-decane exhibits an escaping tendency from the aqueous phase propelling the n-decane molecule into the active site of the CYP. This phenomenon translates into a lower K_m^c constant which represents both ES and substrate-solvent interactions. Additionally, Equation 3.15 based K_m^a constants of CYP153 towards *n*-decane could have been higher compared to other K_m^a constants due to the notable high activity coefficient determined by Equation 3.15 for *n*-decane compared to *n*-octane and *n*-nonane. Therefore, the substitution of *n*-decane's activity coefficient derived from Equation 3.15 into Equation 3.16 ($K_m^a = \gamma^{\infty} \cdot K_m^c / 55.5\ M$), resulted in a high K_m^a value. This suggested that Equation 3.16 may be unsuitable for the determination of K_m^a constants for compounds (determined from solubility) that possess large activity coefficients such as *n*-decane.

In comparison to the K_m^a constants by using activity coefficients from Equation 3.15, the UNIF-LL based K_m^a constants of CYP153A6 showed proportionalities of 57-, 54- and 73-fold lower towards *n*-octane, *n*-nonane and *n*-decane, respectively (Table 4.16). Especially towards the longer chain alkane *n*-decane, this further demonstrated the UNIF-LL model's propensity to underestimate activity coefficients when compared to those determined using Equation 3.15. Because UNIF-LL has a tendency to underestimate activity coefficients, it is likely that K_m^a constants calculated using activity coefficients from Equation 3.15 were more reliable. Both methods for determining activity coefficients demonstrated the same pattern with both enzymes towards *n*-octane and *n*-nonane, showing that the K_m^c towards *n*-octane and *n*-nonane underestimated the enzyme-substrate affinity. For *n*-decane, UNIF-LL based K_m^a constants showed that the enzyme-substrate affinity was underestimated. In contrast, Equation 3.15-based K_m^a constants for the CYPs showed that the K_m^c constants overestimated their affinity towards *n*-decane.

The K_m^a constants for CYP153A6 towards *n*-octane was the highest compared to *n*-nonane and *n*-decane (Table 4.16). This meant that the affinity of CYP163A6 towards *n*-octane was lowest of all the

substrates, as initially shown by K_m^c in Chapter 4.3 and confirmed by the K_m^a constants determined in Table 4.16. While the K_m^c constants indicated CYP153A6 had the highest affinity towards *n*-decane, followed by *n*-nonane and *n*-octane, the K_m^a constants suggested that the enzyme had an affinity towards *n*-nonane that was approximately twice that of *n*-octane and *n*-decane (Table 4.16). The K_m^c constant of CYP153A6 towards *n*-nonane was found to be twice as high as that for *n*-decane, indicating that *n*-decane has a better fit within the active site of CYP153A6 than *n*-nonane. On the other hand, the K_m^a constant of CYP153A6 towards *n*-decane was approximately double that of *n*-nonane. The K_m^a constants in Table 4.16 indicate that *n*-nonane is a superior and better-fitting substrate for the binding pocket of CYP153A6 compared *n*-decane. Therefore, upon excluding decane-solvent and nonane-solvent interactions from the enzyme-substrate interactions (by calculating K_m^a constants), it was clear that the interaction between *n*-nonane and CYP153A6 was more favourable (in terms of shape complementarity and affinity) than the interaction between *n*-decane and CYP153A6.

Despite having a lower hydrophobicity, *n*-nonane's odd-chain structure may have contributed to the higher affinity for the enzyme's active site. These results suggested that 2LPS linear alkane whole cell biohydroxylation may be influenced by the odd-even effect of liquid alkanes. These findings support the results in Chapter 4.3, with K_m^a constants decreasing with an increase in chain length with the exception of *n*-nonane and odd chain liquid alkanes, which introduce deviations and variability into the relationship between hydrophobic substrates and their enzyme-binding affinities (Table 4.16). The odd-even effect of liquid alkanes was extensively discussed in Chapter 4.3.

The analysis also showed that the relationship between K_m^c and substrate hydrophobicity and solubility is not simple. This is because K_m^c constants include the unfavourable molecular interactions that occur between less soluble substrates and aqueous phase molecules, thereby reflecting lower K_m^c constants. These results also suggested that the higher affinity of CYP153A6 towards *n*-decane was due to unfavourable repulsive molecular interactions between *n*-decane and cytoplasmic aqueous molecules, increasing contact between the enzyme and *n*-decane and facilitating its entry into the binding pocket of the active site, reflected by a decrease in the K_m^c constant. Similarly, the higher K_m^c constant of CYP153A6 for *n*-nonane can be attributed to favourable interactions with the surrounding aqueous molecules due to its higher aqueous solubility compared to *n*-decane.

The enzyme's active site is a heme center inside a hydrophobic substrate binding pocket consisting of a few amino acid residues. These residues play a role in substrate orientation and binding (Seifert and Pleiss, 2009, Ji et al., 2013). Longer-chain alkanes have different shapes and sizes and are typically bulkier than shorter chains; therefore, substrate orientation and binding are different for alkanes of different chain lengths. The present study illustrated that odd-chain lengths also display behaviour that impacts substrate binding. The analysis of K_m^a constants in the present chapter deepened the understanding of substrate binding to the hydrophobic pocket of the CYP153 enzymes.

According to the K_m^c constants, CYP153A6 had a 17-fold higher affinity for *n*-decane had than *n*-octane (Table 4.16). The UNIF-LL and Equation 3.15-based K_m^a constants of CYP153A6 revealed that the difference in affinity between *n*-decane and *n*-octane was only 1.5 and 1.13-fold respectively, with a higher affinity towards *n*-decane. The higher affinity (17-fold) depicted by K_m^c constants for CYP153A6 towards *n*-decane was attributed to unfavourable molecular interactions between *n*-decane and aqueous solvent due to its low aqueous solubility, rather than favourable binding affinity between the enzyme and substrate. The K_m^a constants revealed that the true affinities of CYP153A6 towards *n*-octane and *n*-decane were similar, with a slightly higher preference for *n*-decane.

According to the K_m^c constants, CYP153A13 had twice the affinity towards *n*-decane compared to *n*-octane (Table 4.17). In contrast, K_m^a constants obtained using UNIF-LL and Equation 3.15 revealed that CYP153A13 had an affinity for *n*-octane that was six to eight-fold higher, respectively, than that for *n*-decane (Table 4.17). Therefore, the lower K_m^c constant observed for CYP153A13 towards *n*-decane was attributed to unfavourable molecular interactions with cytoplasmic aqueous solvent molecules, rather than favourable affinity between CYP153A13 and *n*-decane. These results demonstrate that *n*-octane is a superior and more suitable substrate for CYP153A13 than *n*-decane.

K_m^a constants can also be compared to substrate solubility as an indicator of whether a reaction is limited by substrate concentration. The MM model predicts that if the substrate solubility is greater than the K_m constants, there are sufficient concentrations of unbound substrate near the enzyme's active site and therefore substrate concentration does not limit the rate of the reaction.

The K_m^c and K_m^a constants of CYP153A6 were lower than the solubilities of the respective substrates, therefore substrate concentration did not limit the reaction rates (Tables 4.16 and 4.17). The K_m^c and K_m^a constants for reactions catalysed by CYP153A13 were higher than the solubilities of the respective substrates, except for the K_m^a constant obtained for *n*-octane using the UNIF-LL model. In this case, the solubility of *n*-octane (4.00 μM) was higher than the K_m^a constant of 0.12763 μM (Table 4.16). This was attributed to potential inaccuracies caused by the underestimation of activity coefficients in the UNIF-LL predictions. Therefore, further investigation was done to determine whether the reaction of CYP153A13 on *n*-octane was limited by substrate concentration in Chapter 4.5, along with the remaining enzyme-substrate pairs.

4.4.3 Concluding remarks

Overall, the results and analyses above demonstrate the applicability of thermodynamic activity in the kinetics of alkane hydroxylation by CYP153 enzymes contained within whole cells. The present study used thermodynamic activity coefficients to reflect the true affinity of each ES pair by taking into account solvent-substrate interactions and excluding the impact of the aqueous cytoplasmic solvent on the substrate. In particular for the primary hydroxylation of alkanes, there is limited research on developing effective methods to investigate the kinetics of whole cell biocatalysis in multiphase systems. The aim was to develop and implement a method for determining the MM kinetic parameters of a 2LPS whole cell linear alkane primary hydroxylation that combine existing techniques such as least squares NLR and activity coefficient-based thermodynamic calculations. As stated previously, it should be noted that the comparisons made were of the general trends between substrates by each enzyme, rather than statistically significant comparisons, due to the error associated with the NLR model predictions made in Chapter 4.3.

Considering even-chain lengths, the results demonstrated that longer chain alkanes exhibit tighter binding to the active site of the CYP153 enzyme, attributed to their higher hydrophobicity and lower aqueous solubility. As previously highlighted, the nature or physicochemical properties of the alkane (related to the odd-even effect) may contribute to pronounced binding for odd-chain lengths such as *n*-nonane, suggesting a potential area for further research.

The results indicate that, assuming that the MM model describes the kinetics of the system sufficiently, the reactions catalysed by CYP153A13 were potentially limited by substrate concentration, while those catalysed by CYP153A6 were not. In a whole cell biocatalytic 2LPS, however, factors other than solubility are involved in regulating the substrate concentration in the bottom aqueous phase and the whole cell aqueous cytoplasm. Thermodynamic phase equilibrium, interfacial LL transfer, and

transmembrane transport all have an impact on the aqueous substrate concentration Therefore, additional research is needed to validate and explore these results to gain a deeper understanding of the underlying mechanisms and reaction rate limitations.

The potential inaccurate estimations of K_m^c taken from the literature for purified enzyme systems and used in Chapter 4.3 may in turn result in inaccurate K_m^a constants. The Equation 3.16 used to calculate K_m^a from K_m^c and activity coefficients was developed by Ferrario et al. (2018) for application in binary water-substrate systems with isolated enzymes, rather than in tertiary water-substrate-cosolvent whole cell systems such as the one investigated in the present study. Certain factors were not taken into account in the analyses, such as cellular health, and potential structural changes in enzymes upon exposure to solvent, substrate, and product. It was assumed that there were no residual effects, including solvent-enzyme interactions or membrane damage caused by solvent molecules or enzyme inhibition (Grosch et al., 2017). The study also assumed that solvent molecules did not interact with the enzyme's active site. In the present study, the chosen solvent molecule, BEHP, is too large to permeate the cellular membrane, validating this assumption (Sandoval et al., 2001).

The impact of different alkane substrate chain length, solubility, and other physicochemical properties, including odd versus even chain length and its corresponding properties were investigated to determine the impact of solvation on biocatalytic reactions. A more accurate evaluation of the ES affinity was made possible by the use of activity coefficients in the aqueous phase. The UNIF-LL process flowsheet model simulations can assist in the determination of activity coefficients, as well as substrate availability in the aqueous phase at thermodynamic equilibrium, covered in the next chapter.

4.5 The impact of aqueous substrate availability on the kinetics of CYP153 whole cell biohydroxylation of C_8, C_9, and C_{10} linear alkanes

As discussed previously, the reaction's velocity, the enzyme-substrate affinity and substrate-aqueous solvent interactions are affected by the length of the alkane chain. In addition, substrate availability is key and also impacts the relevance of the MM model. The objectives of this component of the study were to understand how the reaction kinetics were impacted by the liquid-liquid (LL) phase equilibrium and therefore substrate availability of various alkane chain lengths (C_8, C_9 and C_{10}). Whole cell biohydroxylation performed on C_8, C_9 and C_{10} alkanes was presented in Chapter 4.2 and is used here to assess the potential limitations of the respective substrate partitioning behaviour on the rate of the reactions. The observations made in Chapter 4.5 speak to general trends, rather than statistically significant quantitative assessments due to the lack of formal certainty of the nonlinear regression model predictions, experimental data as well as those by the UNIF-LL model.

The substrate availability in the aqueous phase was determined using the same UNIF-LL process flowsheet model as in Chapter 4.4 (*viz.* Chapter 3.5.1 for the method). The availability and distribution of the substrate in the aqueous phase was investigated to determine whether the biocatalytic reactions were constrained by substrate concentration. It was assumed that the forward reaction catalysed by CYP153 whole cells was favoured over the reverse reaction. Therefore, the chemical equilibrium was assumed to have less significance than the phase equilibrium (Ivanov et al., 1996). By gaining an understanding of alkane partitioning and LL phase equilibrium and how these affect reaction kinetics, improvement of process design, reaction conditions and scalability of whole cell biocatalyic systems can be sought. Further, parameters that act as rate-limiting factors to the biocatalytic productivity can be identified.

The phase equilibria of each of the alkanes were determined using the UNIF-LL predictive process flowsheet model (Banerjee, 1985, Speight, 2018) (Chapter 3.5.1). While in Chapter 4.4, UNIF-LL models allowed the prediction of activity coefficients (Fredenslund et al., 1975), here the UNIF-LL model is used to predict thermodynamic behaviour of biocatalytic 2LPSs, such as liquid-liquid phase equilibrium and alkane partitioning behaviour and between two liquid phases. (Anantpinijwatna et al., 2016, Banerjee, 1985, Brandolín et al., 2022, Lewandowski, 2008, Magnussen et al., 1981). Both experimental conditions and altered conditions, such as the addition of dimethyl sulfoxide (DMSO) or an alternative cosolvent were simulated using the developed UNIF-LL predictive model (described in Chapters 3.5 and 3.6). To gain an understanding of the impact of alkane solubility and associated chain length on their distribution between the organic (Phase I: BEHP-alkane) and aqueous (Phase II) phases, process flowsheet models of the 2LPS were conducted on Aspen Plus (V12) (*viz.* Chapters 3.5 and 3.6 for the method).

4.5.1 The determination of the impact of alkane solubility and chain length on partition coefficients, distribution coefficients, and phase equilibrium constants

The UNIF-LL process flowsheet model was used to predict the mole fractions (x) and activity coefficients (γ^{∞}) of each alkane in the organic phase (α) and aqueous phase (β) at phase equilibrium at 20˚C and 1 atmospheric pressure to simulate the physicochemical properties of the experimental 2-liquid phase system (Figure 4.12). The organic phase is presented in Figure 4.12 as "ORG1" and the aqueous phase as "AQ". "LIQ1" and "LIQ2" represent the two liquid streams simulated at liquid-liquid phase equilibrium (Figure 4.12).

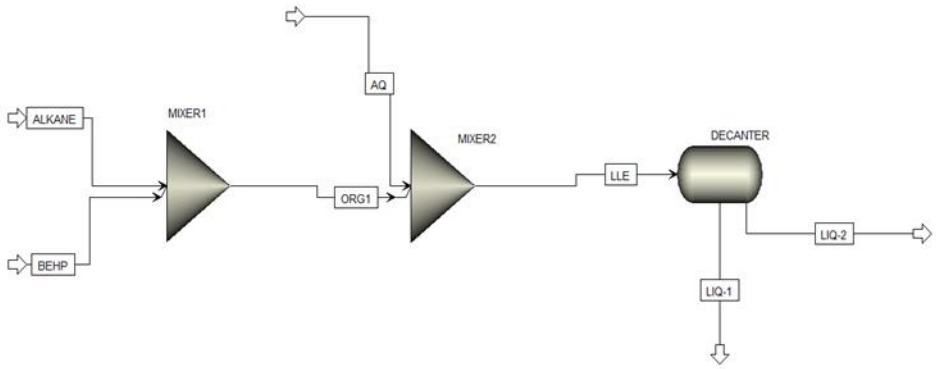


Figure 4.12 Aspen Plus UNIF-LL process flowsheet model of the 2LPS simulating experimental laboratory methods without whole cell biocatalyst. Mole fractions and activity coefficients of alkane in two immiscible phases at minimal Gibbs free energy were provided as an output (LIQ1 and LIQ2). The input stream labelled "ORG1" constitutes mole fractions of the BEHP carrier solvent and alkane (organic phase) previously mixed in MIXER1. The input stream labelled "AQ" constitutes dH2O (aqueous phase). Two streams (ORG1 and AQ) were mixed in MIXER2 and fed into the "DECANTER". The outlet stream labelled "LLE" forms the resultant two-liquid phase stream at phase equilibrium.

Following the determination of mole fractions and activity coefficients, the behaviour of alkanes with different chain lengths with regard to partitioning coefficients (K_p) were calculated (Table 4.18) (Chapter 3.6.1) (Guduru et al., 2019). The K_p, which is measure of the relative affinity of a solute for the organic or aqueous phase in a liquid-liquid system, was determined from the relative ratio of the mole fraction of the solute in the aqueous (x_α) and organic (x_β) phases at equilibrium ($K_p = x_\alpha / x_\beta$) (Chapter 3.6.1, Equation 3.17).

The data relevant to partitioning between phases, generated from the flowsheet model, are presented and compared in Table 4.18. These are compared for each alkane: octane, nonane and decane, at operating conditions used in the experimental studies presented in Chapter 4.2.

As expected, the mole fraction partitioning to the aqueous phase decreased with increasing chain length while the activity coefficient in both phases increased, marginally in the organic phase and by an order of magnitude in the aqueous phase. In accordance, the partition coefficients increased with an increase in alkane chain length, showing an increase in the affinity to the organic phase (Table 4.18). This was attributed to the higher hydrophobicity of longer alkane chain lengths and associated higher affinity for the organic phase rich in highly hydrophobic BEHP. The affinity of the alkanes for the aqueous phase correspondingly decreased. These results were in agreement with findings by Gharbi et al. (2019).

Table 4.18 UNIF-LL predicted mole fractions (x), activity coefficients (γ^{∞}) partition coefficients (K_p), distribution coefficients (K_d) of alkanes in the 2LPS at 20˚ C using a ratio of 1:0.3 aqueous to organic phase and 2:1 ratio of alkane to BEHP for the organic phase assuming infinite dilution

Substrate	*n*-octane	*n*-nonane	*n*-decane
Mole fraction (organic phase) (x_α)	8.29×10^{-1}	8.15×10^{-1}	8.02×10^{-1}
Mole fraction (aqueous phase) (x_β)	35.8×10^{-7}	10.4×10^{-7}	3.03×10^{-7}
Activity coefficient (organic phase) (γ_α)	1.06	1.077	1.08
Activity coefficient (aqueous phase) (γ_β)	2.44×10^5	8.40×10^5	286×10^5
Partition coefficient (K_p)	2.31×10^5	7.86×10^5	264×10^6

Figure 4.13 demonstrates a negatively correlated relationship between alkane chain length and the partitioning determined, showing that the K_p coefficients increased as alkane chain length increased, and its solubility decreased. The partitioning coefficients obtained in the present study were aligned with the $\log P_{oct}$ values (the logarithm of the partition coefficient of the alkane between octanol and water) at 25˚C reported by Abraham et al. (1994) using general linear solvation energy equations using experimentally determined descriptors such as excess molar refraction, solute dipolarity, effective solute hydrogen-bond acidity and basicity and characteristic volume (Abraham et al., 1994).

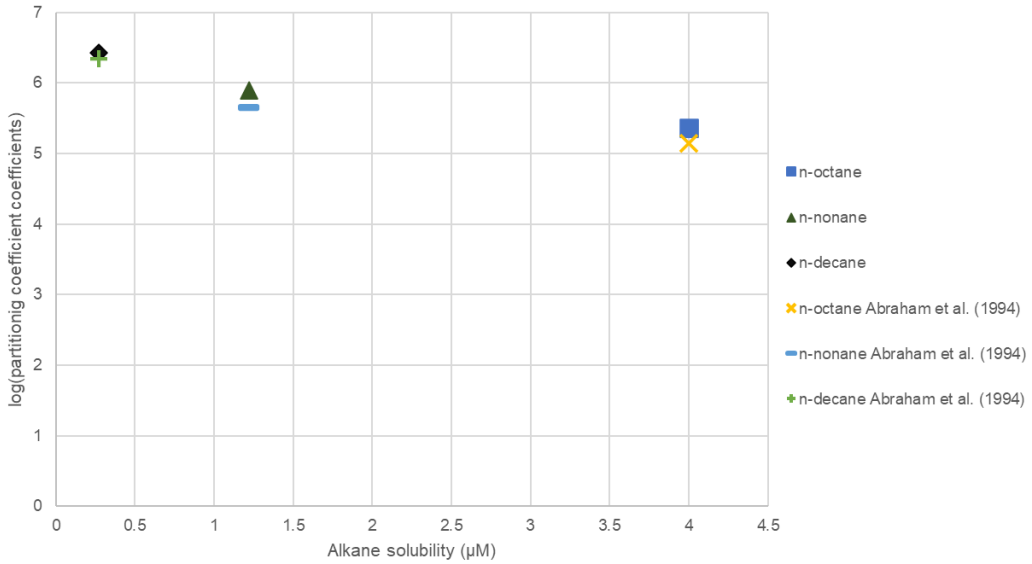


Figure 4.13 The relationship between alkane solubility and UNIF-LL predicted the logarithmic value of the partition coefficients of alkanes in a 2LPS containing aqueous media as the bottom phase and BEHP solvent as the top phase. This was predicted at 20˚ C using a ratio of 1:0.3 aqueous to organic phase and 2:1 ratio of alkane to BEHP for the organic phase assuming infinite dilution. The figure included data reported by Abraham et al. (1994) of the $\log P_{oct}$ values of the alkanes at 25˚C.

4.5.2 Substrate concentration in the aqueous phase

To better understand the effects of substrate availability on the reaction rate and the kinetics of whole cell alkane hydroxylation, the UNIF-LL predicted alkane concentration of c_{alkane}^{aq} in the absence and presence of several cosolvents at equilibrium at 20˚C was determined (Figure 4.14). To investigate the influence of different cosolvents on substrate availability in the aqueous phase, several commonly used and biocompatible cosolvents (BEHP, DMSO, ethyl oleate, hexadecene) were selected (Figure 4.14).

While DMSO, ethyl oleate and hexadecene have been used in previous whole cell biocatalytic systems (Olaofe et al., 2013, Schmid et al., 1998, Kolmar et al., 2019, Foresti et al., 2008), BEHP was specifically chosen to closely resemble and mimic the conditions of the experimental setup employed in the previous chapters of the present study.

To validate the results, the concentration of alkane in the aqueous phase of an alkane-water mixture, in the absence of cosolvent (see 'no cosolvent' in Figure 4.14) were compared to solubility values obtained by Riddick et al. (1986) who used a method of titration to determine the solubility of alkanes in water at 20 ˚C. Riddick et al. (1986) obtained equilibrium compositions of alkane-water mixtures with known masses of alkane, allowing the samples to equilibrate for two hours. Concentrations were determined by measuring the refractive index of the samples, obtaining compositions from a calibration curve, and then determining concentrations from a ternary diagram.

Figure 4.14 illustrated that the 'no cosolvent' UNIF-LL results, which simulated an alkane-water mixture, predicted 1.26, 1.09 and 1.33-fold higher aqueous phase concentrations than the solubility reported by Riddick et al. (1986) for *n*-octane, *n*-nonane and *n*-decane, respectively. It appears that the UNIF-LL model overestimated the solubility of the alkane in the aqueous phase compared to the experimentally determined solubility reported by Riddick et al. (1986). The overestimation was a result of the underestimated activity coefficients predicted by the UNIF-LL model, which was previously discussed and illustrated in Chapter 4.4 (*viz*. Figure 4.11). The underestimation of activity coefficients in the aqueous phase means that the UNIF-LL model assumes a higher degree of ideality or weaker water-alkane interactions than what observed in reality. Therefore, the UNIF-LL model predicted higher solubilities in the aqueous phase than what would be measured experimentally. These findings are supported by findings in studies by Ferrario et al. (2018) and Banerjee et al. (1985) who similarly found that the UNIF-LL model tends to underestimate the activity coefficients of alkanes in water. The explanation of this was discussed in Chapter 4.4, which highlighted that larger hydrophobic compounds possess several chemical functional groups, resulting an accumulated error resulting from the additivity of small deviations associated with each chemical group. This leads to an overall underestimation of activity coefficients and contributes to the observed discrepancies in solubility predictions.

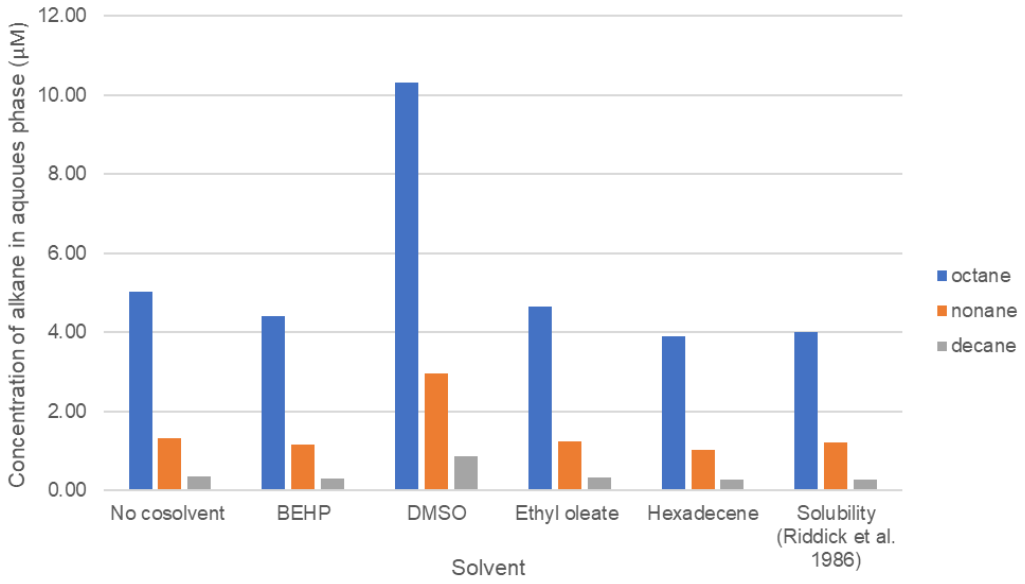


Figure 4.14 UNIF-LL predicted concentration of different linear alkanes in the aqueous phase of a two-liquid-phase system in the absence of solvent (no cosolvent) and presence of cosolvent (BEHP, DMSO, Ethyl Oleate, Hexadecane). UNIF-LL predictions were simulated at 20˚ C at atmospheric pressure. The solubility of alkanes at 20˚C reported by Riddick et al. (1986) is shown at the far right to compare and validate UNIF-LL predicted results.

As expected, the predicted concentration of the alkane in the aqueous phase c_{alkane}^{aq}, using the UNIF-LL approach, decreased as alkane solubility decreased and chain length increased (Figure 4.14). The concentrations of *n*-nonane and *n*-decane in the aqueous were predicted to be 1.16 μM and 0.31 μM, respectively, in the presence of BEHP (Figure 4.14). In comparison, *n*-octane aqueous phase concentration (4.41 μM) were 3.8- and 14.2-fold higher in the presence of BEHP, respectively, and the c_{alkane}^{aq} of *n*-nonane was 3.7-fold higher than *n*-decane (Figure 4.14). The results demonstrated that the developed model using UNIF-LL predicted that the alkane concentration in the aqueous phase decreases as alkane chain length increases, due to the decreased aqueous solubility and increased hydrophobicity, and that these predictions compared well with the values of Reddick et al. (1986).

The c_{alkane}^{aq} of the alkanes predicted in the absence of cosolvent was approximately 1.15-fold higher than the concentration predicted in the presence of BEHP. These findings indicated that the presence of BEHP may further limit the substrate availability during the biocatalytic reaction. Cosolvents play an important role in regulation of substrate and product concentrations to minimise their toxicity towards the whole cell biocatalysts in the aqueous phase. There is a fine balance needed between sufficient substrate availability to improve reaction rates and the need to protect the whole cells from substrate toxicity. These findings are in agreement with Olaofe et al. (2013), who demonstrated that the presence of BEHP improved the rate of product formation by CYP153A6 whole cells compared to the rates achieved by whole cell biocatalysts in the absence BEHP. Olaofe et al. (2013) further demonstrated that the presence of BEHP resulted in a sustained product formation rate over a 120-hour duration, whereas the product formation rate achieved by whole cells in the absence of BEHP levelled off after 48 hours already, indicating decreased cellular health in the absence of BEHP. Although the alkane substrates have higher $\log P_{OW}$ values than their primary alcohol counterparts, indicating lower toxicity, they too exhibit toxicity towards cells (*viz* Table 4.7 in Chapter 4.2).

These results , however, do not demonstrate the ‘cosolvent effect’, which lowers the interfacial tension between the solute and aqueous molecules, lowering the free energy required for the solute to partition into the aqueous phase to improve its solubility (Lewandowski, 2008). The only cosolvent that

demonstrated the 'cosolvent effect' was DMSO, indicated by the increased concentration of alkane in the aqueous phase compared to other solvents (Figure 4.14). Considering the role of the cosolvent effect in increasing substrate solubility, these effects improve alkane availability leading to potential improvements in productivity and reaction rates achieved by whole cell biocatalysts in a 2-liquid-phase-system. For example, the results in Figure 4.14 illustrated that DMSO and ethyl oleate facilitate higher substrate concentrations in the aqueous phase. These results supported evidence in a study by Grant et al. (2012), who showed that substrate access to the aqueous-residing biocatalyst improved by adding DMSO. Although DMSO may improve substrate solubility, it has previously been shown to have detrimental effects on the hydroxylation activity by cytochrome P450 enzymes (Grant et al., 2012). Furthermore, DMSO could create inhibition either due to increasing the aqueous concentration of the product or by having deleterious effects on the whole cell or enzyme itself (Grant et al., 2012). BEHP, on the other hand, has shown to enhance the stability of biocatalysts by reducing the toxic effects of substrate and product, and therefore has led to the selection of BEHP as a cosolvent in systems containing CYP enzymes (Schrewe et al., 2014, Olaofe et al., 2013). However, it is equally important to consider the regulation of concentration of toxic compounds within the aqueous phase.

As previously highlighted, the rate of the biohydroxylation reaction is influenced by the substrate concentration (Schewe et al., 2009) as well as multiple other factors. For example, substrate concentration may be in excess, but if enzyme-substrate affinity is unfavourable, increasing the substrate concentration would not necessarily improve the reaction rate to the desired extent. Further, in a mixture, the enzyme with higher affinity for a particular substrate and substrate with highest affinity for a particular enzyme will access these more effectively. Affinity between enzymes and substrates is an important consideration in determining the impact of substrate concentration on reaction rates.

The Michaelis-Menten (MM) constant K_m indicates the enzyme-substrate (ES) affinity and is a measure of the substrate concentration at which the reaction rate achieves half of its maximum velocity. As stated previously, the thermodynamic activity-based constant K_m^a indicates the true affinity of enzyme towards a substrate. A higher affinity between the enzyme and substrate is indicated by a lower K_m^a constant, i.e., the reaction proceeds at half its maximum velocity at a lower substrate concentration. To determine whether substrate concentration limits the reaction rate achieved by the whole cell biocatalyst, the K_m^a constant (Table 4.21) is compared to the substrate concentration (c_{alkane}^{aq}) (presented in Table 4.20). The UNIF-LL predicted concentration of substrate partitioned ($[c_{alkane}^{aq}]$) into the aqueous phase was used as the substrate concentration ($[S]$) exposed to the whole cell biocatalyst (Table 4.21). For simplicity, from here onwards, c_{alkane}^{aq}, the substrate concentration predicted to be in the aqueous phase at equilibrium, is referred to as $[S]$. The comparison is summarised in Table 4.22 (See footnote [b] of Table 4.22). If the K_m^a is greater than the substrate concentration, the reaction is likely limited by substrate concentration (See footnote [b] in Table 4.22).

Additionally, according to the MM model, the steady-state assumption is valid when the substrate concentration is greater than the enzyme concentration ($[S] \gg [E]$). Therefore, to determine whether the steady-state assumption was valid, $[S]$ (Table 4.20) was compared to the enzyme concentration ($[E]$) (Table 4.19), the latter of which was determined using CO-difference spectral analyses (Chapter 3.2.2). This method of comparison is summarised in Table 4.22 (See footnote [a] of Table 4.22).

The relevant concentrations of enzyme, substrate concentration in the aqueous phase and K_m^a constants, determined in the experimental studies and predicted through the modelling approaches discussed earlier in Chapter 4, were shown in Table 4.19, 4.20 and 4.21, respectively. These were summarised in Table 4.22 by the two comparative methods[a and b].

Table 4.19 Cytochrome P450 concentrations measured using CO-difference spectra in biological repeat 1, 2 and 3 at the start of the reaction (t_0) after 24 hours of IPTG induction at 20˚C and 200 rpm

Biological repeat	CYP153A6	CYP153A13
BR1	0.553	0.227
BR2	0.673	0.175
BR3	0.767	0.307

Table 4.20 Substrate concentrations in the aqueous phase under initial experimental conditions predicted using the UNIF-LL model and process flowsheet

Alkane	Substrate concentration in the aqueous phase ($[S]$) (μM)
n-octane	4.41
n-nonane	1.16
n-decane	0.33

Table 4.21 Thermodynamic activity-based Michaelis-Menten constants (K_m^a) of enzyme-substrate pairs determined using UNIF-LL activity coefficients.

Enzyme	Alkane	UNIF-LL-based K_m^a (μM)
CYP153A6	*n*-octane	0.000748
	n-nonane	0.0003328
	n-decane	0.0005149
CYP153A13	*n*-octane	0.1276
	n-nonane	n.d
	n-decane	0.8238

The K_m^a constants in Table 4.21 were determined in Chapter 4.4.2.

Table 4.22. A table summarising the results of two methods used to identify the validity of the steady state assumption[a] and whether enzyme-substrate reactions that are limited by substrate[b]

Enzyme-substrate pair	Is $[S] \ll [E]$? [a]	Is $K_m^a \gg [S]$? [b]
CYP153A6-octane	No	No
CYP153A6-nonane	No	No
CYP153A6-decane	Yes	No
CYP153A13-octane	No	No
CYP153A13-nonane	No	n.d[c]
CYP153A13-decane	No	Yes

[a] If substrate concentration is smaller than the enzyme concentration (indicated by yes), then the steady-state assumption is not valid
[b] If the K_m^a is greater than the substrate concentration (indicated by yes) then the reaction is dependent on substrate concentration and limited by substrate concentration
[c] The K_m^a constant of CYP153A13 towards *n*-nonane was not determined because an initial estimate was not available for nonlinear regression analysis in Chapter 4.3

In BR1, BR2 and BR3, the concentration of CYP153A6 was $[E]_0$ =0.553 μM, 0.673 μM and 0.767 μM respectively (Table 4.19), which were all lower than the predicted aqueous phase concentrations for *n*-octane (4.4113 μM) and *n*-nonane (1.1618 μM), respectively (Table 4.20). Therefore, the steady-state assumption was valid for the CYP153 whole cell biohydroxylation of *n*-octane and *n*-nonane because $[S] \gg [E]$ (Table 4.22). The concentration of *n*-octane ([S]= 4.4113 µM) in the aqueous phase was greater than the K_m^a constant determined for CYP153A6 (UNIF-LL -K_m^a = 0.000748 μM) (Tables 4.20 and 21). These findings illustrated that the alkane hydroxylation rate by CYP153A6 on *n*-octane was independent of substrate concentration (Table 4.22). Similarly, the concentration of *n*-nonane (1.1618 μM) was greater than the K_m^a constant (UNIF-LL K_m^a = 0.000333). Overall, both methods of comparison[a] and [b] revealed that the steady-state assumption was valid for CYP153A6 biohydroxylation of *n*-octane and *n*-nonane and that these reaction rates were not limited by the substrate concentration. These findings further suggest that factors such as pH, temperature, enzyme concentration, enzyme-substate affinity, ineffective electron coupling, or enzymatic intrinsic reaction mechanisms may be limiting to the reaction rate achieved by CYP153A6 for the hydroxylation of *n*-octane and *n*-nonane.

In contrast, the steady-state assumption was invalid for *n*-decane because the concentration of *n*-decane, $[S]$ =0.3113 μM was lower than the measured concentration of CYP153A6, $[E]_0$ =0.673 and 0.767 μM in BR2 and BR3. The K_m^a constant of CYP153A6 towards *n*-decane, was 0.000515 μM (UNIF-LL based). The *n*-decane concentration was 0.3113 μM , greater than the K_m^a constant. By comparing K_m^a constants to $[S]$[b], it was illustrated that the rate of reaction catalysed by CYP153A6 whole cells towards *n*-decane was limited by substrate concentration.

CYP153A13 had a measured concentration of $[E]_0 = 0.227\ \mu M$ in BR1, 0.175 μM in BR2 and 0.307 μM in BR3, which were lower than the concentrations of *n*-octane and *n*-nonane, 4.4113 and 1.1618 μM, respectively. The predicted *n*-decane concentration, $[S]$ =0.3113 μM, was higher than the

concentration of CYP153A13, $[E]_0$ =0.175 μM in BR2. Therefore, for the first method of comparison[a], $[S] \gg [E]$ for CYP153A13 catalysed reactions on *n*-octane, *n*-nonane, and *n*-decane (Table 4.22). This illustrated that the steady state assumption was valid for the rate of the reactions catalysed by CYP153A13 on all the alkanes tested. By comparing the K_m^a constants to the substrate concentration, it was determined whether reaction rates catalysed by CYP153A13 were limited by substrate concentration (Table 4.22). The UNIF-LL based K_m^a constant of CYP153A13 towards *n*-octane was 0.1276 μM (Table 4.21), which was lower than the *n*-octane's aqueous phase concentration of 4.4113 μM (Table 4.20). Therefore, the comparison[b] illustrated that substrate concentration was not limiting to the reaction rate catalysed by CYP153A13 for the primary hydroxylation of *n*-octane (Table 4.22). This implied that the reaction rates by CYP153A13 were limited by other factors such as low enzyme concentrations, poor electron coupling, ineffective enzymatic reaction mechanisms or electron coupling, or even poor affinity between enzyme and substrate.

As stated previously, the comparison[a] of enzyme to substrate concentration showing that the *n*-decane concentration was greater than CYP153A13's concentration indicating that the steady state model was appropriate for analysis of the kinetics of CYP153A13 towards n-decane. The results of the second comparison of K_m^c constants to $[S]$ demonstrated that substrate limitation was likely (Table 4.22). The UNIF-LL based K_m^a constant obtained for CYP153A13 towards *n*-decane was 0.8238 μM (Table 4.21) which was greater than the $[S]$ of 0.3113 μM (Table 4.20), indicating that the reaction rate was limited by substrate concentration (Table 4.22). Therefore, the evidence suggests that the hydroxylation rate of *n*-decane by CYP153A13 was limited by substrate concentration. Additionally, in Chapter 4.1, strong evidence showed that CYP153A13 exhibits very low expression levels, so it is likely that both substrate and enzyme concentrations in combination contributed to a lack primary alcohol production altogether.

It was demonstrated in Chapter 4.3 that CYP153A6 performs better on *n*-decane than on *n*-nonane in terms of maximum TOF and reaction velocities. The reasons for this were unclear; the limited interfacial liquid-liquid or transmembrane transport of *n*-nonane compared to *n*-decane was postulated. Here, this result is further investigated. The alkane *n*-decane is 5-fold lower in aqueous solubility than *n*-nonane. In contrast, in Chapter 4.4, it was shown that CYP153A6 has nearly double the affinity for *n*-nonane over *n*-decane. Here, it was shown that the concentration of *n*-nonane was not limiting to CYP153A6's reaction rate, whereas for the reaction rate towards *n*-decane was limited by substrate concentration (Table 4.22).

This leaves room for two possible explanations. Firstly, *n*-decane may have physicochemical properties that favour more effective transmembrane transport compared to *n*-nonane. This could be due to the 'odd-even' effect of liquid alkanes, discussed through the study. Secondly, the relevant toxicity of *n*-nonane compared to *n*-decane towards the whole cells may drive the behaviour. A commonly used indicator of toxicity towards cellular health is the $\log P_{OW}$ of the compound given in Table 4.7. According to $\log P_{OW}$ values, *n*-nonane ($\log P_{OW}$ = 5.45) is more toxic to the *E. coli* recombinant cells than *n*-decane ($\log P_{OW}$ = 5.98). Therefore, it is possible that CYP153A6 whole cells achieved lower TOFs on *n*-nonane because the cellular health was more compromised than in systems containing *n*-decane. Overall, substrate concentration was confirmed to not be rate-limiting for the biohydroxylation of CYP153A6 on *n*-nonane, but lower TOFs were achieved on *n*-nonane than *n*-decane. This means that other factors limited the rate of the reaction, such as transmembrane transport or compound toxicity were rate-limiting.

4.5.1 Concluding remarks

It is simple to assume that the reaction rates in biocatalytic 2LPSs are primarily constrained by the substrate concentration exposed to the biocatalyst due to the limited solubility of alkanes in water, but the results show that this may not always be the case and many factors should be considered when drawing such a conclusion. The analyses showed the importance of considering additional factors such as enzyme concentration, substrate concentration available in the aqueous phase and enzyme-substrate affinity when determining whether the reaction is constrained by substrate concentration.

The findings showed that, despite the MM model's widespread application in enzyme kinetic analyses, its steady state assumption may not always hold true in certain situations where substrate partitioning across the liquid-liquid interface is limited, leading to a substrate concentration that is lower than the enzyme concentration. Furthermore, the rates of some ES pair reactions were determined to be independent and not limited by substrate concentration because their MM constants were lower than the aqueous phase substrate concentration at phase equilibrium. The present study paved the way for a deeper comprehension of the effects of phase equilibrium, substrate partitioning behaviour and substrate availability on the reaction rates of the biohydroxylation of medium chain alkanes catalysed by CYP153 enzymes.

These findings relate to the thermodynamic properties of the initial state of the 2LPS, and how these properties impact the reaction rates and kinetics of CYP153 whole cell bio hydroxylation of medium-chain alkanes. It is important to note that the state of the 2LPS was simulated using predictive UNIF-LL modelling without a whole cell biocatalyst. In the presence of a biocatalyst, upon ES contact, substrate reacts to form the desired primary alcohol product, resulting in substrate depletion and product accumulation. As a result, the chemical gradient across the LL interface changes because of the differential solvation of the substrate and product, establishing successive and multiple equilibrium states, which was not accounted for in the UNIF-LL process flowsheet model. This illustrates the demand-based delivery nature of the whole cell biocatalytic 2LPS. Further, the impact of oxygen availability was not studied directly in the 2LPSs under investigation, owing to this effect having been previously shown not to impact results in the reaction system used. It is noted that the use of different alkane substrates may lead to different oxygen solubilities in the aqueous media and therefore differing oxygen availability for the biohydroxylation reaction. This could have potentially impacted the kinetics of the reaction. However, the findings by Meissner et al. (2013) who showed that oxygen was not limiting to biotransformation performance of CYP153A6 whole cell alkane oxidation suggested that this effect was not a critical variable.

5 Conclusions and Recommendations

5.1 Introduction

The overall goal of the study was to examine the applicability of a multi-component CYP153A6 and a fusion CYP153 cytochrome P450 enzyme system in the specific biocatalysis of medium chain linear alkanes (C_8 – C_{10}) to primary alcohols and to determine how solvation and substrate-solvent interactions influenced the kinetics of whole cell biohydroxylation of alkanes in a two-liquid-phase-system, with potential for applicability to other similar biocatalysis reactions. Alkane chain length plays a significant role in determining the properties of the final product, impacting its market value and scope of use in the final chemical sector. It is noteworthy that the biohydroxylation of odd-chain alkanes (such as *n*-nonane) using CYP153 whole cell biocatalysts appears to have received little attention in past literature. The cytochrome P450 enzymes, which are well known for successful hydroxylation of medium chain alkanes, were selected as candidates to evaluate the effect of carbon length, substrate solubility, and solvation effects, on parameters such as the turnover frequency, maximum reaction velocity, and Michaelis-Menten (MM) constant K_m.

The study compared two highly specific enzymes of the CYP153 family: CYP153A6 is a multi-component enzyme system and CYP153A13 is a fusion enzyme. The expression of these two enzymes was confirmed and compared using protein characterisation techniques. These activity of these two enzymes was tested on three substrates: *n*-octane, *n*-nonane and *n*-decane. This was achieved by using whole cell biocatalysts in a 2-liquid-phase-system.

It was hypothesised that alkane accessibility to the enzyme in the whole cells significantly constrained the biocatalytic activity in 2-liquid phase systems, because of the transfer of alkane across the aqueous-organic interface and cellular membrane. To test this hypothesis, in vivo biohydroxylation tests on alkanes of different chain lengths (C_8, C_9, C_{10}) which have different solubilities were performed. Whole cells were placed in an aqueous phase and alkanes were applied to the organic phase, which formed a two-liquid-phase-system (2LPS). Primary alcohol product concentration was measured over time.

This experimental approach allowed investigation into the substrate scope for two CYP153 enzymes (CYP153A6 and CYP153A13) and the impacts of substrate solubility on the turnover frequency (TOF), reaction rate and kinetics. Alkanes of different chain length possess different solubilities in the aqueous phase of the 2LPS and this in turn impacts substrate availability to the whole cell biocatalyst. Furthermore, alkanes of different chain length also show different toxicity to whole cells, indicated by their $\log P_{OW}$ values. Moreover, the extent to which alkanes partition into the aqueous phase also contributes the amount that the whole cells are exposed to the substrate which decreases cellular health. This highlights the major trade-off between the concentration that maintains cellular health and substrate availability to the biocatalyst sufficient for desirable high reaction rates. Therefore, by analysing product concentration and TOFs achieved by the whole cell biocatalysts over time, conclusions could be drawn on the impact of toxicity on the reaction rates and performance of each ES pair.

To investigate the potential lack of substrate availability to the enzymes and its impact on the kinetics of the reaction, factors such as phase equilibrium, partition coefficient, enzyme-substrate affinity, enzyme concentration and substrate-solvent effects were considered. These factors enabled clarification and placed into context whether the rate of the reaction of a particular enzyme on a substrate was limited by substrate availability. Enzyme-substrate affinity is an important consideration

in determining the impact of substrate availability on the reaction rate. Excess substrate may be available in the surrounds of an enzyme's active site, but if the enzyme has a low affinity for the substrate, increasing the concentration of substrate will have a lower impact on the reaction rate than with a high affinity.

For the determination of enzyme-substrate affinity, both the apparent and true enzyme Michaelis-Menten constants were determined. The apparent MM constants (K_m^c) were determined using the experimentally determined approach mentioned above. Using this approach, the product concentrations as a function of time were fitted to the Michaelis-Menten model using least squares nonlinear regression. Apparent kinetics allowed insight into the overall performance by determination of maximum reaction velocities (v_{max}) and K_m^c constants of the two enzyme systems on C_8, C_9 and C_{10} alkanes, respectively.

The resultant apparent K_m^c constants, however, include variables such as substrate-solvent interactions, transmembrane transport and assumes that the enzyme is exposed to an excess substrate concentration. Therefore, it was considered that the interaction of solvent molecules with substrate impacts the effective concentration of substrate available to the enzyme's active site. This made it necessary to determine true MM constants based on thermodynamic activity (K_m^a), which exclude substrate-solvent interactions, to account only for the effective substrate concentration accessible to the whole cell biocatalyst. K_m^a constants were determined using activity coefficients (γ) of each alkane in the aqueous phase at 20˚C. Activity coefficients were determined by developing a process flowsheet and modelling it using the UNIF-LL model on Aspen Plus (V12), as well as by a theoretical method using Equation 3.15 ($c^{sol} = 55.5\ M/\gamma^{\infty}$) relating activity coefficient, substrate solubility and the mole-based concentration of pure liquid water at standard state (55.5 M). The process flowsheet model simulated the conditions of experimental 2-liquid-phase-system at 20˚C, 1 atmospheric pressure, using a ratio of aqueous to organic phase (containing BEHP) in the absence of the whole cell biocatalyst.

By following this approach, the true affinity between each enzyme and substrate pair was revealed. This area of biocatalysis has not been extensively studied, especially on CYP153-containing whole cells for primary alkane oxidation in 2LPSs The K_m^a constants assist in determining the appropriate concentration of substrate for a reaction (Tipton et al., 2014). Therefore, accurate determination of K_m^a constant facilitates process design of biocatalytic systems. Additionally, the enzyme-substrate (ES) pair with the lowest K_m is favoured in biocatalytic systems with multiple enzymes or limited supply of various substrates in a 'one-pot system'- a system that is desirable for large scale industrial application.

The UNIF-LL process flowsheet model was also used to predict the concentration of alkane concentration in the aqueous phase, alkane partitioning and phase equilibrium behaviour of the 2LPS. By selecting alkanes of different solubilities, the understanding of substrate solubility on phase equilibrium and partitioning behaviour of alkane into the aqueous phase in a 2LPS was enhanced. In turn, the reactions in which the steady state assumption of the Michaelis-Menten constant held true were identified by comparing enzyme concentration to substrate concentration. Additionally, the rates of ES reaction pairs that were indeed limited by substrate concentration were identified by comparing the K_m^a constants to substrate concentration.

It has been commonly assumed that substrate solubility must be a rate-limiting factor in whole cell biocatalytic 2LPSs, due to its low solubility in the aqueous phase and limiting transport across the cellular membrane and liquid-liquid interface. However, as mentioned previously, factors such as enzyme-substrate affinity, enzyme concentration and substrate partitioning into the aqueous phase are

very important to consider in determining whether the biocatalytic reaction is limited by substrate supply. In the present study, these factors were holistically considered to identify which of the alkane hydroxylation reactions catalysed by CYP153 whole cells in the 2-liquid-phase system were truly limited by substrate concentration.

In summary, the results demonstrated that whole cell biocatalytic 2LPSs are complex, and that productivity relies on a number of interdependent factors. Overall, the present study integrated aspects of molecular biology and biochemistry, such as enzyme expression in recombinant host organisms and analyses of enzyme-substrate affinity, into the context of chemical engineering science, such as thermodynamics and solvation effects. This integrative approach is crucial for developing and advancement of biocatalytic processes that are applicable to industry.

5.2 Outcomes of study

The CYP153A6 and CYP153A13 enzymes were recombinantly expressed in the *E. coli* BL21(BE3) host organism under IPTG induction for 24 hours at 20˚C and thereafter, their expression was characterised. CO-difference spectral analyses revealed that CYP153A6 concentration (0.0859 ± 0.0089 $\mu mol_{P450} \cdot g_{DCW}^{-1}$) was double that of CYP153A13 (0.0420 ± 0.0065 $\mu mol_{P450} \cdot g_{DCW}^{-1}$). Moreover, SDS-PAGE gels illustrated that CYP153A13 experienced some degree of unwanted cytoplasmic proteolysis and cleavage of the peptide linker between the CYP and reductase domain of the fusion enzyme.

Despite the challenges of CYP153A13 expression, *in vivo* biohydroxylation tests using low cell density cultures revealed that CYP153A13 achieved a comparable turnover number (TON) to CYP153A6 after 24 h reaction. This illustrated that while the CYP153A6 concentrations were higher in the bioreaction mixture, which allowed more frequent contact between enzyme and substrate to facilitate higher bioconversion, the benefits of superior electron coupling within the CYP153A13 fusion system likely contributed to its ability in achieving a similar TON to CYP153A6. By 48 hours, the TON achieved by CYP153A6 was three times higher than achieved by CYP153A13. At this point of the reaction, substrate is less available than at 24 h, because it has been consumed. CYP153A13 has lower expression than CYP153A6 and therefore CYP153A13 cultures would experience a decrease in enzyme-substrate contact in reaction mixtures containing less substrate than CYP153A6 which is more abundantly available for enzyme-substrate contact. These results may also indicate that CYP153A6 has a higher affinity towards *n*-octane than CYP153A13.

Further *in vivo* biohydroxylation tests were conducted on *n*-octane, *n*-nonane and *n*-decane in a 2-liquid-phase-system to investigate the substrate scope of CYP153A6 and CYP153A13 whole cells and the impact of substrate solubility on achieved TOFs as the reaction proceeds. The product concentration, measured over time, was used to calculate the turnover frequency (TOF) as a function of time. This investigation revealed a few key findings. Firstly, alkanes that have lower solubility showed later onset of initiation of bioconversion. Reactions catalysed by CYP153A13 too showed later onset of reaction initiation than CYP153A6. This revealed that substrate solubility and enzyme concentration both play a key role in the time taken for a reaction to start.

The TOF achieved by CYP153A6 was the highest on *n*-octane, followed by *n*-decane with the lowest TOF achieved on *n*-nonane. The TOF achieved by CYP153A13 was the highest on *n*-octane followed by *n*-nonane with no activity on *n*-decane. From these findings it could be concluded that an increase in substrate solubility results in higher TOFs achieved by the CYP153A6 and CYP153A13 whole cell biocatalysts due to an increase in substrate available to contact the enzyme's active site. Interestingly however, the odd-even effect of liquid alkanes may have resulted in preferential transmembrane

transport of *n*-decane compared to *n*-nonane, explaining why CYP153A6 had achieved higher TOFs on *n*-decane. This was related to the ability and preference of bacterial species in nature to degrade even-chain alkanes.

The profiles of the TOF as a function of time revealed that CYP153A6 showed a more rapid increase of TOFs over time to reach the TOF maximum compared to CYP153A13, illustrating the impact of lower CYP153A13 concentration, or/and CYP153A6's higher affinity towards the respective alkane substrates. CYP153A13 showed no activity on *n*-decane as a result of low enzyme and substrate concentration levels. Also, both CYP153A6 and CYP153A13-biocatalysed reactions of alkanes with higher solubility showed a more rapid increase in TOFs over time, which resulted from higher available substrate to the active site of the enzyme leading to more frequent contact between the substrate and enzyme. This is because alkanes with higher solubility were more readily transported across the liquid-liquid interface and cellular membrane.

Although higher substrate availability may have increased the TOFs, a trade-off between substrate concentration and toxicity was demonstrated for TOF as a function of time. The relative toxicity of the alkanes was illustrated by the decrease in TOF achieved by CYP153A6 whole cells from 24 to 48 hours on *n*-octane, while an increase was observed for *n*-decane from 24 to 48 hours. Octane has a lower $\log P_{OW}$ value than *n*-decane, indicating that octane has higher toxicity than *n*-decane. Additionally, the increased partitioning of *n*-octane into the aqueous phase contributes to higher substrate exposure to the cells, resulting in biocatalyst instability and detrimental effects to cellular health. Moreover, 1-octanol has the highest toxicity compared to the other primary alcohols, as indicated by its low $\log P_{OW}$ value of 2.42. As substrate is consumed, product is formed and therefore also is toxic to cellular health. Therefore, CYP153A6 showed a decreased in TOF on *n*-octane due to the formation its relatively toxic primary alcohol product. Since 1-decanol has a higher $\log P_{OW}$ value of 4, it is less toxic to the cells than 1-octanol, which was indicated by the increased in TOFs by CYP153A6 from 24 to 48 h.

Overall, it was shown that to improve CYP153A6 biocatalytic systems, strategies to maintain cellular health should be implemented, in addition to the BEHP carrier solvent that was used to regulate substrate and product concentration, and thereby exposure to the whole cell catalyst, and has previously been shown to enhance biocatalyst stability. In contrast, CYP153A13 biocatalytic systems should be improved by increasing the expression, and resultant stable concentration, of enzyme to improve frequency of contact with substrate.

The above experimental studies were used to determine the apparent Michaelis-Menten parameters: the maximum reaction velocity v_{max} and the Michaelis-Menten (MM) constant K_m^c. Apparent kinetic analyses were conducted by fitting product as a function of time data to the MM model equation using least squares nonlinear regression. The apparent v_{max} determined for the CYP153A6 and CYP153A13 whole cell biohydroxylation of alkanes showed similar results to the TOF analyses above, showing higher maximum reaction velocities were achieved by CYP153A6 than CYP153A13. For both whole cell biocatalysts, reactions performed on alkanes with higher solubility achieved higher maximum reaction velocities. Also, the maximum reaction velocity of CYP153A6 was highest for *n*-octane at 0.008 $\mu mol \cdot L_{BRM}^{-1} \cdot min^{-1}$, followed by *n*-decane and lastly *n*-nonane.

The K_m^c is commonly used as indicator of enzyme-substrate affinity and represents the substrate concentration at which half the v_{max} is achieved. The apparent K_m^c constants determined revealed that as alkane chain length increases, the affinity between the enzyme and substrate correspondingly increases. This was the case for both enzymes on all three alkanes tested. The K_m^c constant however, reflects both enzyme-substrate and substrate-solvent interactions and assumes excess substrate was

exposed to the enzyme's active site. Substrate-solvent molecular interactions impact the substrate's availability to access the active site of the enzyme. To exclude the impact of solvent interactions with the substrate and reveal true affinity between enzyme and substrate, thermodynamic activity-based MM constants (K_m^a) of each ES pair were determined. The thermodynamic activity of the substrate in the aqueous phase was determined by developing a UNIF-LL process flowsheet model.

While K_m^c values depicted that CYP153A6 had the highest affinity for *n*-decane, K_m^a constants revealed that CYP153A6 has a higher affinity to *n*-nonane than *n*-decane. Therefore, the K_m^c constants obtained were an indication of unfavourable molecular interactions between *n*-decane and an aqueous solvent rather than favourable ES affinity. The K_m^a constants revealed that CYP153A6 had the lowest affinity towards *n*-octane, and the highest affinity towards *n*-nonane. For CYP153A13, the K_m^c showed double the affinity to *n*-decane than *n*-octane while K_m^a values showed a six-fold higher affinity for *n*-octane. Therefore the K_m^c constant of CYP153A13 was lower towards *n*-decane because of unfavourable molecular interactions between *n*-decane and water rather than favourable ES affinity.

This approach was necessary for the study of future systems that contain more than one enzyme in 'one-pot' systems, and in systems that lack substrate supply, which will result in favourable biocatalysis for the ES pair with the lowest K_m. Moreover, the determination of true K_m^a constants can facilitate the selection of the suitable substrate concentration range to be applied in a biocatalytic system. These insights pave way for the advancement and development of larger scale biocataytic systems relevant to industry.

As stated previously, enzyme-substrate affinity is a crucial factor to consider while determining whether a reaction is limited by substrate supply. The ES pairs that were in fact limited by substrate availability were identified by comparing the enzyme concentration and K_m^aconstants to the UNIF-LL predicted substrate concentration in the aqueous phase. The comparisons revealed that the hydroxylation of *n*-decane catalysed by CYP153A13 whole cells was likely to be limited by substrate concentration and lack of substrate accessibility, whereas the reactions catalysed by CYP153A6 towards *n*-octane and *n*-nonane were not. Together with the findings that CYP153A13 showed low expression and partial truncation, the reaction catalysed by CYP153A13 was limited by both substrate and enzyme concentration.

Overall, the findings paved the way for perspectives and analyses of whole cell biocatalytic kinetics in 2LPSs. In turn, this could assist in scaling-up and process control (Bolivar and Nidetzky, 2013). The investigations confirmed that, for some enzyme-substrate pairs, the reaction in 2LPSs is limited by the relatively small alkane concentrations present in the aqueous phase. This limits techno-economic feasibility and imparts high operational and downstream processing costs (Puetz et al., 2020) From a process operation perspective, these findings can inform the substrate concentration range to be applied and which reactions would be favoured under those conditions. In summary, the importance of thermodynamic analysis of substrate partitioning and decoupling substrate-solvent interactions from ES interactions to reveal true affinities of ES pairs was highlighted, thereby providing insight for 'one-pot' with multiple enzymes and limited substrate supply. Moreover, the importance of functional enzyme expression was highlighted as a factor that could be majorly limiting to biocatalytic systems.

5.3 Recommendations

The findings of this study could guide the selection of strategies to improve the performance of systems under study. For example, since it was found that the reaction catalysed by CYP153A6 on *n*-octane and *n*-nonane were not limited by substrate concentration, further optimisation may involve improving

the reaction mechanism by enzyme engineering techniques or using cell-free enzyme preparations to eliminate transport limitations while ensuring these enzyme preparations are resistant to the detrimental impacts of solvents, thereby improving the lifespan and stability of the whole cell catalyst. To improve the activity achieved by the CYP153A6 system, strategies to maintain cellular health over time are recommended. CYP153A6 showed the highest affinity towards *n*-nonane, yet the lowest TOFs were achieved on *n*-nonane, suggesting transmembrane transport limitation. The findings supported a strong recommendation to investigate the odd-even effect of liquid alkanes in biocatalytic two-liquid-phase systems. For example, it would be beneficial to perform the experimental 2LPS with a different cosolvent to understand the effect of BEHP on odd-even chain lengths more deeply, especially in the case on *n*-nonane.

The variability observed between different biological repeats BR1, BR2 and BR3 should be minimised, by repeating more experiments or potentially by using glycerol stocks that are well-maintained to avoid the introduction of genetic heterogeneity amongst the subpopulations. In the present study, the comparison of different substrates could only be done within each biological repeat, rather than across biological repeats. Ideally, the biological repeats should have been consistent and used together to enhance the accuracy and statistical significance of the observations made.

Since CYP153A13 showed cytoplasmic proteolytic cleavage and low expression levels, it is clear that the expression system needed improvement; potential approaches include using different promotors or co-expression of molecular chaperones (Choi et al., 2006). For CYP153A13, enzyme engineering techniques to limit and prevent the cytoplasmic cleavage of CYP153A13 are recommended. Moreover, since the results showed that the reaction catalysed by CYP153A13 was limited by *n*-decane concentration, it is recommended to increase the substrate concentration by adding surfactants, cosolvent screening or even the use of micro-reactors.

Typically, kinetic parameters are measured by collecting experimental data from a series of reactions individually starting with an initial substrate concentration. In the present study, a method was developed to determine the kinetic parameters at a single substrate concentration which is governed by phase equilibrium and substrate partitioning into the aqueous phase. Phase equilibrium represents a single thermodynamic state, which has ongoing variability as substrate is consumed and product is formed. The UNIF-LL process flowsheet model only simulated the initial state of the 2LPS, in the absence of product. Therefore, mass transfer analysis is recommended to provide rates of substrate transfer across the liquid-liquid interface. Kinetic analysis of the reaction rate within the whole cell could then be nested into the mass transfer model to obtain a holistic view of the kinetic rates of the biocatalytic reactions. Modifications to the UNIF-LL model would also improve the accuracy of the analysis because of its under-estimation of activity coefficients, supported by findings of previous studies. COSMO-RS models, which are based on computational quantum mechanics, have been shown to provide highly accurate estimates of activity coefficients for biocatalytic 2LPSs and are recommended for future study (Grosch et al., 2017, Navas et al., 2010, Klamt, 2005, Gerber and Soares, 2010).

6 References

ABRAHAM, M. H., ACREE JR, W. E. & ZISSIMOS, A. M. 2017. The correlation and prediction of infinite dilution activity coefficients of compounds in water at 298.15 K. *Fluid Phase Equilibria,* 449**,** 117-129.

ABRAHAM, M. H., CHADHA, H. S., WHITING, G. S. & MITCHELL, R. C. 1994. Hydrogen bonding. 32. An analysis of water-octanol and water-alkane partitioning and the Δlog P parameter of Seiler. *Journal of pharmaceutical sciences,* 83**,** 1085-1100.

ABRAMS, D. S. & PRAUSNITZ, J. M. 1975. Statistical thermodynamics of liquid mixtures: a new expression for the excess Gibbs energy of partly or completely miscible systems. *AIChE Journal,* 21**,** 116-128.

ABU, R. & WOODLEY, J. M. 2015. Application of enzyme coupling reactions to shift thermodynamically limited biocatalytic reactions. *ChemCatChem,* 7**,** 3094-3105.

AL-HAQUE, N., SANTACOLOMA, P. A., NETO, W., TUFVESSON, P., GANI, R. & WOODLEY, J. M. 2012. A robust methodology for kinetic model parameter estimation for biocatalytic reactions. *Biotechnology Progress,* 28**,** 1186-1196.

ANANTPINIJWATNA, A., KIM, S. H., SALES-CRUZ, M., O'CONNELL, J. P. & GANI, R. 2016. Application of the e-KT-UNIFAC Model for the Improved and Innovative Design of Biphasic Reacting Systems. *Journal of Chemical & Engineering Data,* 61**,** 4090-4103.

ARBUCKLE, W. B. 1983. Estimating activity coefficients for use in calculating environmental parameters. *Environmental Science & Technology,* 17**,** 537-542.

AYALA, M. & TORRES, E. 2004. Enzymatic activation of alkanes: constraints and prospective. *Applied Catalysis A: General,* 272**,** 1-13.

BAKKES, P. J., RIEHM, J. L., SAGADIN, T., RÜHLMANN, A., SCHUBERT, P., BIEMANN, S., GIRHARD, M., HUTTER, M. C., BERNHARDT, R. & URLACHER, V. B. 2017. Engineering of versatile redox partner fusions that support monooxygenase activity of functionally diverse cytochrome P450s. *Scientific Reports,* 7**,** 1-13.

BANERJEE, S. 1985. Calculation of water solubility of organic compounds with UNIFAC-derived parameters. *Environmental Science & Technology,* 19**,** 369-370.

BARBERIS, S., QUIROGA, E., MORCELLE, S., PRIOLO, N. & LUCO, J. M. 2006. Study of phytoproteases stability in aqueous-organic biphasic systems using linear free energy relationships. *Journal of Molecular Catalysis B: Enzymatic,* 38**,** 95-103.

BELL, S. G., ORTON, E., BOYD, H., STEVENSON, J.-A., RIDDLE, A., CAMPBELL, S. & WONG, L.-L. 2003. Engineering cytochrome P450cam into an alkane hydroxylase. *Dalton Transactions***,** 2133-2140.

BEZERRA, R. M. & DIAS, A. A. 2007. Utilization of integrated Michaelis-Menten equation to determine kinetic constants. *Biochemistry and Molecular Biology Education,* 35**,** 145-150.

BLANK, L. M., IONIDIS, G., EBERT, B. E., BÜHLER, B. & SCHMID, A. 2008. Metabolic response of Pseudomonas putida during redox biocatalysis in the presence of a second octanol phase. *The FEBS journal,* 275**,** 5173-5190.

BOLIVAR, J. M. & NIDETZKY, B. 2013. Multiphase biotransformations in microstructured reactors: opportunities for biocatalytic process intensification and smart flow processing. *Green Processing and Synthesis,* 2**,** 541-559.

BOMMARIUS, A. S. & PAYE, M. F. 2013. Stabilizing biocatalysts. *Chemical Society Reviews,* 42**,** 6534-6565.

BORDEAUX, M., DE GIRVAL, D., RULLAUD, R., SUBILEAU, M., DUBREUCQ, E. & DRONE, J. 2014. High-cell-density cultivation of recombinant Escherichia coli, purification and characterization of a self-sufficient biosynthetic octane ω-hydroxylase. *Applied Microbiology and Biotechnology,* 98**,** 6275-6283.

BORDEAUX, M., GALARNEAU, A. & DRONE, J. 2012. Catalytic, mild, and selective oxyfunctionalization of linear alkanes: current challenges. *Angewandte Chemie International Edition,* 51**,** 10712-10723.

BORDEAUX, M., GALARNEAU, A., FAJULA, F. & DRONE, J. 2011. A regioselective biocatalyst for alkane activation under mild conditions. *Angewandte Chemie International Edition,* 50**,** 2075-2079.

BRAIUCA, P., EBERT, C., BASSO, A., LINDA, P. & GARDOSSI, L. 2006. Computational methods to rationalize experimental strategies in biocatalysis. *TRENDS in Biotechnology,* 24**,** 419-425.

BRANDOLÍN, S. E., SCILIPOTI, J. A., ANDREATTA, A. E. & MAGARIO, I. 2022. UNIFAC Evaluation of the Liquid–Liquid Phase Equilibrium during Lipase-Catalyzed Peracidation of Different Carboxylic Acids. *Journal of Chemical & Engineering Data,* 67**,** 1994-2003.

BRAUN, S., IMRE, A. & KRASKA, T. 2013. Stability limits of n-nonane calculated from molecular dynamics interface simulations. *The Journal of Chemical Physics,* 138**,** 244710.

BRINK, L., TRAMPER, J., LUYBEN, K. C. A. & VAN'T RIET, K. 1988. Biocatalysis in organic media. *Enzyme and Microbial Technology,* 10**,** 736-743.

BÜHLER, B., BOLLHALDER, I., HAUER, B., WITHOLT, B. & SCHMID, A. 2003. Use of the two-liquid phase concept to exploit kinetically controlled multistep biocatalysis. *Biotechnology and Bioengineering,* 81**,** 683-694.

CAVANI, F. & TRIFIRO, F. 1997. Some aspects that affect the selective oxidation of paraffins. *Catalysis Today,* 36**,** 431-439.

CELEBI, A. T., JAMALI, S. H., BARDOW, A., VLUGT, T. J. & MOULTOS, O. A. 2020. Finite-size effects of diffusion coefficients computed from molecular dynamics: a review of what we have learned so far. *Molecular Simulation***,** 1-15.

CHEN, L.-Q. 2019. Chemical potential and Gibbs free energy. *MRS Bulletin,* 44**,** 520-523.

CHOI, B., REMPALA, G. A. & KIM, J. K. 2017. Beyond the Michaelis-Menten equation: Accurate and efficient estimation of enzyme kinetic parameters. *Scientific Reports,* 7**,** 1-11.

CHOI, J. H., KEUM, K. C. & LEE, S. Y. 2006. Production of recombinant proteins by high cell density culture of Escherichia coli. *Chemical Engineering Science,* 61**,** 876-885.

CIARAMELLA, A., MINERDI, D. & GILARDI, G. 2017. Catalytically self-sufficient cytochromes P450 for green production of fine chemicals. *Rendiconti Lincei,* 28**,** 169-181.

COOLEY, R. B., DUBBELS, B. L., SAYAVEDRA-SOTO, L. A., BOTTOMLEY, P. J. & ARP, D. J. 2009. Kinetic characterization of the soluble butane monooxygenase from Thauera butanivorans, formerly 'Pseudomonas butanovora'. *Microbiology,* 155**,** 2086.

CORNELISSEN, S., JULSING, M. K., VOLMER, J., RIECHERT, O., SCHMID, A. & BÜHLER, B. 2013. Whole-cell-based CYP153A6-catalyzed (S)-limonene hydroxylation efficiency depends on host background and profits from monoterpene uptake via AlkL. *Biotechnology and Bioengineering,* 110**,** 1282-1292.

CRABTREE, R. H. & LEI, A. 2017. Introduction: CH activation. ACS Publications.

DELVIGNE, F. & GOFFIN, P. 2014. Microbial heterogeneity affects bioprocess robustness: Dynamic single-cell analysis contributes to understanding of microbial populations. *Biotechnology Journal,* 9**,** 61-72.

DHIMAN, I., BERG, M. C., PETRIDIS, L., SMITH, J. C. & GAUTAM, S. 2022. Odd-even effect in n-alkane systems: A molecular dynamics study. *Physical Chemistry Chemical Physics*.

DOMAŃSKA, U., OKUNIEWSKI, M., OKUNIEWSKA, P., PADUSZYŃSKI, K. & TUROWSKI, T. 2015. Phase equilibrium and bioproduction of the aroma compound 2-phenylethanol in a biphasic aqueous system. *European Food Research and Technology,* 240**,** 1177-1186.

DONOVAN, R. S., ROBINSON, C. W. & GLICK, B. 1996. Optimizing inducer and culture conditions for expression of foreign proteins under the control of thelac promoter. *Journal of Industrial Microbiology,* 16**,** 145-154.

DUETZ, W. A., VAN BEILEN, J. B. & WITHOLT, B. 2001. Using proteins in their natural environment: potential and limitations of microbial whole-cell hydroxylations in applied biocatalysis. *Current Opinion in Biotechnology,* 12**,** 419-425.

DUGGLEBY, R. G. 1995. Analysis of enzyme progress curves by nonlinear regression. *Methods in Enzymology.* Elsevier.

ECKSTEIN, M., LEMBRECHT, J., SCHUMACHER, J., EBERHARD, W., SPIESS, A., PETERS, M., ROOSEN, C., GREINER, L., LEITNER, W. & KRAGL, U. 2006a. Maximise equilibrium conversion in biphasic catalysed reactions: How to obtain reliable data for equilibrium constants? *Advanced Synthesis & Catalysis,* 348**,** 1597-1604.

ECKSTEIN, M. F., PETERS, M., LEMBRECHT, J., SPIESS, A. C. & GREINER, L. 2006b. Maximise equilibrium conversion in biphasic catalysed reactions: mathematical description and practical guideline. *Advanced Synthesis & Catalysis,* 348**,** 1591-1596.

EGGERS, D. K., BLANCH, H. W. & PRAUSNITZ, J. M. 1989. Extractive catalysis: solvent effects on equilibria of enzymatic reactions in two-phase systems. *Enzyme and Microbial Technology,* 11**,** 84-89.

EIBEN, S., KAYSSER, L., MAURER, S., KÜHNEL, K., URLACHER, V. B. & SCHMID, R. D. 2006. Preparative use of isolated CYP102 monooxygenases—a critical appraisal. *Journal of Biotechnology,* 124**,** 662-669.

FAM, H. & DAUGULIS, A. J. 2012. Substrate mass transport in two-phase partitioning bioreactors employing liquid and solid non-aqueous phases. *Bioprocess and Biosystems Engineering,* 35**,** 1367-1374.

FENTON, H. 1894. LXXIII.—Oxidation of tartaric acid in presence of iron. *Journal of the Chemical Society, Transactions,* 65**,** 899-910.

FERGUSON, A. L., DEBENEDETTI, P. G. & PANAGIOTOPOULOS, A. Z. 2009. Solubility and molecular conformations of n-alkane chains in water. *The Journal of Physical Chemistry B,* 113**,** 6405-6414.

FERRARIO, V., HANSEN, N. & PLEISS, J. 2018. Interpretation of cytochrome P450 monooxygenase kinetics by modeling of thermodynamic activity. *Journal of Inorganic Biochemistry,* 183**,** 172-178.

FERRARIO, V. & PLEISS, J. 2019. Molecular simulations of enzymes under non-natural conditions. *The European Physical Journal Special Topics,* 227**,** 1631-1638.

FINDEISEN, A., THUM, O. & ANSORGE-SCHUMACHER, M. B. 2014. Biocatalytically active silCoat-composites entrapping viable Escherichia coli. *Applied Microbiology and Biotechnology,* 98**,** 1557-1566.

FORESTI, M., PEDERNERA, M., FERREIRA, M. & BUCALÁ, V. 2008. Kinetic modeling of enzymatic ethyl oleate synthesis carried out in biphasic systems. *Applied Catalysis A: General,* 334**,** 65-72.

FRANCIS JR, R. T. & BECKER, R. R. 1984. Specific indication of hemoproteins in polyacrylamide gels using a double-staining process. *Analytical Biochemistry,* 136**,** 509-514.

FREDENSLUND, A., JONES, R. L. & PRAUSNITZ, J. M. 1975. Group-contribution estimation of activity coefficients in nonideal liquid mixtures. *AIChE Journal,* 21**,** 1086-1099.

FUJITA, N., SUMISA, F., SHINDO, K., KABUMOTO, H., ARISAWA, A., IKENAGA, H. & MISAWA, N. 2009. Comparison of two vectors for functional expression of a bacterial cytochrome P450 gene in Escherichia coli using CYP153 genes. *Bioscience, Biotechnology, and Biochemistry***,** 0907091551-0907091551.

FUNHOFF, E. G., BAUER, U., GARCÍA-RUBIO, I., WITHOLT, B. & VAN BEILEN, J. B. 2006. CYP153A6, a soluble P450 oxygenase catalyzing terminal-alkane hydroxylation. *Journal of Bacteriology,* 188**,** 5220-5227.

FUNHOFF, E. G., SALZMANN, J., BAUER, U., WITHOLT, B. & VAN BEILEN, J. B. 2007. Hydroxylation and epoxidation reactions catalyzed by CYP153 enzymes. *Enzyme and Microbial Technology,* 40**,** 806-812.

FUNHOFF, E. G. & VAN BEILEN, J. B. 2007. Alkane activation by P450 oxygenases. *Biocatalysis and Biotransformation,* 25**,** 186-193.

GARDOSSI, L., POULSEN, P. B., BALLESTEROS, A., HULT, K., ŠVEDAS, V. K., VASIĆ-RAČKI, Đ., CARREA, G., MAGNUSSON, A., SCHMID, A. & WOHLGEMUTH, R. 2010. Guidelines for reporting of biocatalytic reactions. *Trends in Biotechnology,* 28**,** 171-180.

GARIKIPATI, S. J., MCIVER, A. M. & PEEPLES, T. L. 2009. Whole-cell biocatalysis for 1-naphthol production in liquid-liquid biphasic systems. *Appl. Environ. Microbiol.,* 75**,** 6545-6552.

GERBER, R. P. & SOARES, R. D. P. 2010. Prediction of infinite-dilution activity coefficients using UNIFAC and COSMO-SAC variants. *Industrial & Engineering Chemistry Research,* 49**,** 7488-7496.

GLAZYRINA, J., MATERNE, E.-M., DREHER, T., STORM, D., JUNNE, S., ADAMS, T., GRELLER, G. & NEUBAUER, P. 2010. High cell density cultivation and recombinant protein production with Escherichia coli in a rocking-motion-type bioreactor. *Microbial Cell Factories,* 9**,** 42.

GOLIČNIK, M. 2013. The integrated Michaelis-Menten rate equation: déjà vu or vu jàdé? *Journal of Enzyme Inhibition and Medicinal Chemistry,* 28**,** 879-893.

GRANT, C., DA SILVA DAMAS PINTO, A. C., LUI, H. P., WOODLEY, J. M. & BAGANZ, F. 2012. Tools for characterizing the whole-cell bio-oxidation of alkanes at microscale. *Biotechnology and Bioengineering,* 109**,** 2179-2189.

GRANT, C., DESZCZ, D., WEI, Y.-C., MARTINEZ-TORRES, R. J., MORRIS, P., FOLLIARD, T., SREENIVASAN, R., WARD, J., DALBY, P. & WOODLEY, J. M. 2014. Identification and use of an alkane transporter plug-in for applications in biocatalysis and whole-cell biosensing of alkanes. *Scientific Reports,* 4**,** 1-9.

GRANT, C., WOODLEY, J. M. & BAGANZ, F. 2011. Whole-cell bio-oxidation of n-dodecane using the alkane hydroxylase system of P. putida GPo1 expressed in E. coli. *Enzyme and Microbial Technology,* 48**,** 480-486.

GROSCH, J. H., WAGNER, D., NISTELKAS, V. & SPIEß, A. C. 2017. Thermodynamic activity-based intrinsic enzyme kinetic sheds light on enzyme–solvent interactions. *Biotechnology Progress,* 33**,** 96-103.

GUDIMINCHI, R. K., RANDALL, C., OPPERMAN, D. J., OLAOFE, O. A., HARRISON, S. T., ALBERTYN, J. & SMIT, M. S. 2012. Whole-cell hydroxylation of n-octane by Escherichia coli strains expressing the CYP153A6 operon. *Applied Microbiology and Biotechnology,* 96**,** 1507-1516.

GUDURU, S., SINGH, M. L. & PAHARI, P. 2019. Thermodynamic distribution ratio model for PUREX major solutes. *Separation Science and Technology,* 54**,** 1422-1429.

GUENGERICH, F. P., MARTIN, M. V., SOHL, C. D. & CHENG, Q. 2009. Measurement of cytochrome P450 and NADPH–cytochrome P450 reductase. *Nature Protocols,* 4**,** 1245.

HALLING, P., FITZPATRICK, P. F., RAUSHEL, F. M., ROHWER, J., SCHNELL, S., WITTIG, U., WOHLGEMUTH, R. & KETTNER, C. 2018. An empirical analysis of enzyme function reporting for experimental reproducibility: Missing/incomplete information in published papers. *Biophysical Chemistry,* 242**,** 22-27.

HALLING, P. J. 1990. Solvent selection for biocatalysis in mainly organic systems: predictions of effects on equilibrium position. *Biotechnology and Bioengineering,* 35**,** 691-701.

HALLING, P. J. 1994. Thermodynamic predictions for biocatalysis in nonconventional media: theory, tests, and recommendations for experimental design and analysis. *Enzyme and Microbial Technology,* 16**,** 178-206.

HALLING, P. J. 2004. What can we learn by studying enzymes in non–aqueous media? *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences,* 359**,** 1287-1297.

HARVEY, D. 2020. *Liquid-Liquid Extractions* [Online]. : LibreTexts libraries. Available: https://chem.libretexts.org/Courses/BethuneCookman_University/B-CU%3A_CH-345_Quantitative_Analysis/Book%3A_Analytical_Chemistry_2.1_(Harvey)/07%3A_Obtaining_and_Preparing_Samples_for_Analysis/7.07%3A_Liquid-

Liquid_Extractions#:~:text=Partition%20Coefficients%20and%20Distribution%20Ratios,-As%20we%20learned&text=The%20partition%20coefficient%20is%20a,of%20A%20and%20B%20change. [Accessed].

HE, J., ZHANG, S., ZHANG, X., QIAN, Y., HE, H. & WU, H. 2016. Composition and distribution characteristics and geochemical significance of n-alkanes in core sediments in the northern part of the South Yellow Sea. *Journal of Chemistry,* 2016.

HEINEMANN, M., KÜMMEL, A., GIESEN, R., ANSORGE-SCHUMACHER, M. B. & BÜCHS, J. 2003. Experimental and theoretical analysis of phase equilibria in a two-phase system used for biocatalytic esterifications. *Biocatalysis and Biotransformation,* 21**,** 115-121.

HONDA MALCA, S. 2013. *Substrate characterization and protein engineering of bacterial cytochrome P450 monooxygenases for the bio-based synthesis of omega-hydroxylated aliphatic compounds.* University of Stuttgart.

HRISTOVA, M., DONCHEV, T., KOLEV, D., BALOCH, I. & GEORGIEV, H. 2015. Parameter's estimate in Wilson equation. *International Electronic Journal of Pure and Applied Mathematics,* 9.

HUA, F. & WANG, H. Q. 2014. Uptake and trans-membrane transport of petroleum hydrocarbons by microorganisms. *Biotechnology & Biotechnological Equipment,* 28**,** 165-175.

IVANOV, I. P., KASCHE, V. & PETKOV, D. D. 1996. Thermodynamics of enzymic peptide synthesis in biphasic aqueous–organic systems: coupling of the chemical and phase equilibria. *Biocatalysis and Biotransformation,* 14**,** 181-194.

JACOBS, C. L., DO AIDO-MACHADO, R., TOLMIE, C., SMIT, M. S. & OPPERMAN, D. J. 2022. CYP153A71 from Alcanivorax dieselolei: Oxidation beyond Monoterminal Hydroxylation of n-Alkanes. *Catalysts,* 12**,** 1213.

JAKOB, A., GRENSEMANN, H., LOHMANN, J. & GMEHLING, J. 2006. Further development of modified UNIFAC (Dortmund): Revision and extension 5. *Industrial & Engineering Chemistry Research,* 45**,** 7924-7933.

JANSSEN, A. E., VAN DER PADT, A. & RIET, K. V. T. 1995. Prediction of the reaction equilibrium of biocatalytic reactions in aqueous-organic two-phase systems. *Biocatalysis and Biotransformation,* 12**,** 223-240.

JERIČEVIĆ, Ž. & KUŠTER, Ž. 2005. Non-linear optimization of parameters in Michaelis-Menten kinetics. *Croat. Chem. Acta,* 78**,** 519-523.

JI, Y., MAO, G., WANG, Y. & BARTLAM, M. 2013. Structural insights into diversity and n-alkane biodegradation mechanisms of alkane hydroxylases. *Frontiers in Microbiology,* 4**,** 58.

JULSING, M. K., SCHREWE, M., CORNELISSEN, S., HERMANN, I., SCHMID, A. & BÜHLER, B. 2012. Outer membrane protein AlkL boosts biocatalytic oxyfunctionalization of hydrophobic substrates in Escherichia coli. *Applied and Environmental Microbiology,* 78**,** 5724-5733.

JUNG, E., PARK, B. G., AHSAN, M. M., KIM, J., YUN, H., CHOI, K.-Y. & KIM, B.-G. 2016. Production of ω-hydroxy palmitic acid using CYP153A35 and comparison of cytochrome P450 electron transfer system in vivo. *Applied microbiology and biotechnology,* 100**,** 10375-10384.

KAMPERS, L. F., VOLKERS, R. J. & MARTINS DOS SANTOS, V. A. 2019. Pseudomonas putida KT 2440 is HV 1 certified, not GRAS. *Microbial biotechnology,* 12**,** 845-848.

KARA, S. & LIESE, A. 2019. Process Considerations for the Application of Enzymes. *Industrial Enzyme Applications***,** 71-94.

KEIM, W. 2003. Multiphase catalysis and its potential in catalytic processes: the story of biphasic homogeneous catalysis. *Green Chemistry,* 5**,** 105-111.

KEMMER, G. & KELLER, S. 2010. Nonlinear least-squares data fitting in Excel spreadsheets. *Nature protocols,* 5**,** 267-281.

KERTES, A. 1989. IUPAC Solubility Data Series. *by DG Shaw, Pergamon, Oxford,* 37.

KIM, P. Y., POLLARD, D. J. & WOODLEY, J. M. 2007. Substrate supply for effective biocatalysis. *Biotechnology Progress,* 23**,** 74-82.

KLAMT, A. 2005. *COSMO-RS: from quantum chemistry to fluid phase thermodynamics and drug design*, Elsevier.

KOCH, D. J., CHEN, M. M., VAN BEILEN, J. B. & ARNOLD, F. H. 2009a. In vivo evolution of butane oxidation by terminal alkane hydroxylases AlkB and CYP153A6. *Appl. Environ. Microbiol.,* 75**,** 337-344.

KOCH, D. J., CHEN, M. M., VAN BEILEN, J. B. & ARNOLD, F. H. 2009b. In vivo evolution of butane oxidation by terminal alkane hydroxylases AlkB and CYP153A6. *Applied and Environmental Microbiology,* 75**,** 337-344.

KOCHIUS, S., VAN MARWIJK, J., EBRECHT, A., OPPERMAN, D. & SMIT, M. 2018. Deconstruction of the CYP153A6 Alkane Hydroxylase System: Limitations and Optimization of In Vitro Alkane Hydroxylation. *Catalysts,* 8**,** 531.

KOEBMANN, B. J., WESTERHOFF, H. V., SNOEP, J. L., NILSSON, D. & JENSEN, P. R. 2002. The glycolytic flux in Escherichia coli is controlled by the demand for ATP. *Journal of bacteriology,* 184**,** 3909-3916.

KOLMAR, J. F., THUM, O. & BAGANZ, F. 2019. Improving Product Specificity of Whole-Cell Alkane Oxidation in Nonconventional Media: A Multivariate Analysis Approach. *Biotechnology journal,* 14**,** 1800581.

KOVÁROVÁ-KOVAR, K. & EGLI, T. 1998. Growth kinetics of suspended microbial cells: from single-substrate-controlled growth to mixed-substrate kinetics. *Microbiol. Mol. Biol. Rev.,* 62**,** 646-666.

KOZUCH, S. & MARTIN, J. M. 2012. “Turning over” definitions in catalytic cycles. ACS Publications.

KUBOTA, M., NODATE, M., YASUMOTO-HIROSE, M., UCHIYAMA, T., KAGAMI, O., SHIZURI, Y. & MISAWA, N. 2005. Isolation and functional analysis of cytochrome P450 CYP153A genes from various environments. *Bioscience, Biotechnology, and Biochemistry,* 69**,** 2421-2430.

LAANE, C., BOEREN, S., VOS, K. & VEEGER, C. 2009. Rules for optimization of biocatalysis in organic solvents. *Biotechnology and bioengineering,* 102**,** 2-8.

LABINGER, J. A. 2016. Platinum-catalyzed C–H functionalization. *Chemical reviews,* 117**,** 8483-8496.

LABINGER, J. A. & BERCAW, J. E. 2002. Understanding and exploiting C–H bond activation. *Nature,* 417**,** 507-514.

LAUCHNOR, E. G., TOPP, D., PARKER, A. & GERLACH, R. 2015. Whole cell kinetics of ureolysis by S porosarcina pasteurii. *Journal of applied microbiology,* 118**,** 1321-1332.

LAVALLIE, E. R. & MCCOY, J. M. 1995. Gene fusion expression systems in Escherichia coli. *Current opinion in biotechnology,* 6**,** 501-506.

LEE, C. W., YU, S.-C., LEE, J.-H., PARK, S.-H., PARK, H., OH, T.-J. & LEE, J. H. 2016. Crystal structure of a putative cytochrome P450 alkane hydroxylase (CYP153D17) from Sphingomonas sp. PAMC 26605 and its conformational substrate binding. *International Journal of Molecular Sciences,* 17**,** 2067.

LEE, S. Y. 1996. High cell-density culture of Escherichia coli. *Trends in biotechnology,* 14**,** 98-105.

LENNEN, R. M. & PFLEGER, B. F. 2013. Microbial production of fatty acid-derived fuels and chemicals. *Current opinion in biotechnology,* 24**,** 1044-1053.

LEON, R., FERNANDES, P., PINHEIRO, H. & CABRAL, J. 1998. Whole-cell biocatalysis in organic media. *Enzyme and Microbial Technology,* 23**,** 483-500.

LEWANDOWSKI, J. A. 2008. Modified UNIFAC-LLE Group-Interaction Parameters for the Prediction of Gasoline-Ethanol-Water Equilibria.

LIEBERMAN, R. L. & ROSENZWEIG, A. C. 2004. Biological methane oxidation: regulation, biochemistry, and active site structure of particulate methane monooxygenase. *Critical reviews in biochemistry and molecular biology,* 39**,** 147-164.

LUNDEMO, M. T. & WOODLEY, J. M. 2015. Guidelines for development and implementation of biocatalytic P450 processes. *Applied Microbiology and Biotechnology,* 99**,** 2465-2483.

MAGNUSSEN, T., RASMUSSEN, P. & FREDENSLUND, A. 1981. UNIFAC parameter table for prediction of liquid-liquid equilibriums. *Industrial & Engineering Chemistry Process Design and Development,* 20**,** 331-339.

MAIER, T., FÖRSTER, H.-H., ASPERGER, O. & HAHN, U. 2001. Molecular characterization of the 56-kDa CYP153 from Acinetobacter sp. EB104. *Biochemical and biophysical research communications,* 286**,** 652-658.

MAKRIDES, S. C. 1996. Strategies for achieving high-level expression of genes in Escherichia coli. *Microbiol. Mol. Biol. Rev.,* 60**,** 512-538.

MARRERO, J. & GANI, R. 2002. Group-contribution-based estimation of octanol/water partition coefficient and aqueous solubility. *Industrial & Engineering Chemistry Research,* 41**,** 6623-6633.

MAVROVOUNIOTIS, M. L. 1990. Group contributions for estimating standard Gibbs energies of formation of biochemical compounds in aqueous solution. *Biotechnology and Bioengineering,* 36**,** 1070-1082.

MAVROVOUNIOTIS, M. L. 1991. Estimation of standard Gibbs energy changes of biotransformations. *Journal of Biological Chemistry,* 266**,** 14440-14445.

MEINHOLD, P., PETERS, M. W., HARTWICK, A., HERNANDEZ, A. R. & ARNOLD, F. H. 2006. Engineering cytochrome P450 BM3 for terminal alkane hydroxylation. *Advanced Synthesis & Catalysis,* 348**,** 763-772.

MEISSNER, M. P. 2013. *Reactor design for the hydroxylation of n-octane using a cyp153a6-based biocatalytic system expressed in Escherichia coli.* University of Cape Town.

MEYER, D., BU, B. & SCHMID, A. 2006. Process and catalyst design objectives for specific redox biocatalysis. *Advances in applied microbiology,* 59**,** 53-91.

MUNRO, A. W., GIRVAN, H. M. & MCLEAN, K. J. 2007. Cytochrome P450–redox partner fusion enzymes. *Biochimica et Biophysica Acta (BBA)-General Subjects,* 1770**,** 345-359.

NAGAR, S., ARGIKAR, U. A. & TWEEDIE, D. J. 2014. Enzyme kinetics in drug metabolism: fundamentals and applications. *Enzyme Kinetics in Drug Metabolism.* Springer.

NAKANISHI, K., KIMURA, Y. & MATSUNO, R. 1986. Kinetics and equilibrium of enzymatic synthesis of peptides in aqueous/organic biphasic systems: Thermolysin-catalyzed synthesis of N-(benzyloxycarbonyl)-l-aspartyl-l-phenylalanine methyl ester. *European Journal of biochemistry,* 161**,** 541-549.

NAKANISHI, K. & MATSUNO, R. 1986. Kinetics of enzymatic synthesis of peptides in aqueous/organic biphasic systems: Thermolysin-catalyzed synthesis of N-(benzyloxycarbonyl)-l-phenylalanyl-l-phenylalanine methyl ester. *European journal of biochemistry,* 161**,** 533-540.

NARAYANAN, K. 2004. *A textbook of chemical engineering thermodynamics*, PHI Learning Pvt. Ltd.

NAVAS, A., ORTEGA, J., MARTÍN, T. S. & PALOMAR, J. 2010. Thermodynamic Analysis of Systems Formed by Alkyl Esters with α, ω-Alkyl Dibromides: New Experimental Information and the Use of a Dense Database to Describe Their Behavior Using the UNIFAC Group Contribution Method and the COSMO-RS Methodology. *Industrial & Engineering Chemistry Research,* 49**,** 12726-12739.

NEKRASOV, P., PIVEN, O., NEKRASOV, O., GUDZ, O. & KRYVONIS, N. 2018. Kinetics and thermodynamics of biocatalytic glycerolysis of triacylglycerols enriched with omega-3 polyunsaturated fatty acids.

NI, Y. & CHEN, R. R. 2004. Accelerating whole-cell biocatalysis by reducing outer membrane permeability barrier. *Biotechnology and bioengineering,* 87**,** 804-811.

NIE, Y., CHI, C.-Q., FANG, H., LIANG, J.-L., LU, S.-L., LAI, G.-L., TANG, Y.-Q. & WU, X.-L. 2014a. Diverse alkane hydroxylase genes in microorganisms and environments. *Scientific Reports,* 4**,** 4968.

NIE, Y., LIANG, J.-L., FANG, H., TANG, Y.-Q. & WU, X.-L. 2014b. Characterization of a CYP153 alkane hydroxylase gene in a Gram-positive Dietzia sp. DQ12-45-1b and its “team role” with alkW1 in alkane degradation. *Applied Microbiology and Biotechnology,* 98**,** 163-173.

NODATE, M., KUBOTA, M. & MISAWA, N. 2006. Functional expression system for cytochrome P450 genes using the reductase domain of self-sufficient P450RhF from Rhodococcus sp. NCIMB 9784. *Applied microbiology and biotechnology,* 71**,** 455-462.

NORDHOLT, N., VAN HEERDEN, J., KORT, R. & BRUGGEMAN, F. J. 2017. Effects of growth rate and promoter activity on single-cell protein expression. *Scientific reports,* 7**,** 1-11.

O'REILLY, E., KÖHLER, V., FLITSCH, S. L. & TURNER, N. J. 2011. Cytochromes P450 as useful biocatalysts: addressing the limitations. *Chemical Communications,* 47**,** 2490-2501.

OLAOFE, O. A., FENNER, C. J., GUDIMINCHI, R., SMIT, M. S. & HARRISON, S. T. 2013. The influence of microbial physiology on biocatalyst activity and efficiency in the terminal hydroxylation of n-octane using Escherichia coli expressing the alkane hydroxylase, CYP153A6. *Microbial cell factories,* 12**,** 8.

ORTIZ DE MONTELLANO, P. R. 2009. Hydrocarbon hydroxylation by cytochrome P450 enzymes. *Chemical reviews,* 110**,** 932-948.

PALMER, T. & BONNER, P. L. 2007. *Enzymes: biochemistry, biotechnology, clinical chemistry*, Elsevier.

PARK, C. 2022. Visual Interpretation of the Meaning of k cat/KM in Enzyme Kinetics. *Journal of Chemical Education,* 99**,** 2556-2562.

PENNEC, A., JACOBS, C. L., OPPERMAN, D. J. & SMIT, M. S. 2015. Revisiting cytochrome P450-mediated oxyfunctionalization of linear and cyclic alkanes. *Advanced Synthesis & Catalysis,* 357**,** 118-130.

PESCI, L. 2017. *Biocatalytic (de) carboxylation of phenolic compounds: fundamentals and applications.* Technische Universität Hamburg.

PFRUENDER, H., JONES, R. & WEUSTER-BOTZ, D. 2006. Water immiscible ionic liquids as solvents for whole cell biocatalysis. *Journal of Biotechnology,* 124**,** 182-190.

PLEISS, J. 2017. Thermodynamic activity-based interpretation of enzyme kinetics. *Trends in biotechnology,* 35**,** 379-382.

PRPICH, G. P. & DAUGULIS, A. J. 2004. Polymer development for enhanced delivery of phenol in a solid-liquid two-phase partitioning bioreactor. *Biotechnology progress,* 20**,** 1725-1732.

PUETZ, H., PUCHĽOVÁ, E., VRANKOVÁ, K. & HOLLMANN, F. 2020. Biocatalytic Oxidation of Alcohols. *Catalysts,* 10**,** 952.

REZAI, S. A. S. & TRAA, Y. 2008. Equilibrium shift in membrane reactors: A thermodynamic analysis of the dehydrogenative conversion of alkanes. *Journal of Membrane Science,* 319**,** 279-285.

RIDDICK, J. A., BUNGER, W. B. & SAKANO, T. K. 1986. Organic solvents: physical properties and methods of purification.

ROBINSON, P. K. 2015. Enzymes: principles and biotechnological applications. *Essays in biochemistry,* 59**,** 1-41.

ROJAS, A., DUQUE, E., SCHMID, A., HURTADO, A., RAMOS, J.-L. & SEGURA, A. 2004. Biotransformation in double-phase systems: physiological responses of Pseudomonas putida DOT-T1E to a double phase made of aliphatic alcohols and biosynthesis of substituted catechols. *Applied and Environmental Microbiology,* 70**,** 3637-3643.

ROJO, F. 2005. Specificity at the end of the tunnel: understanding substrate length discrimination by the AlkB alkane hydroxylase. *Journal of Bacteriology,* 187**,** 19-22.

RON, E. Z. & ROSENBERG, E. 2002. Biosurfactants and oil bioremediation. *Current opinion in biotechnology,* 13**,** 249-252.

SALEHMIN, M., ANNUAR, M. & CHISTI, Y. 2013. High cell density fed-batch fermentations for lipase production: feeding strategies and oxygen transfer. *Bioprocess and biosystems engineering,* 36**,** 1527-1543.

SANDOVAL, G. C., MARTY, A. & CONDORET, J. S. 2001. Thermodynamic activity-based enzyme kinetics: Efficient tool for nonaqueous enzymology. *AIChE journal,* 47**,** 718-726.

SCHELLER, U., ZIMMER, T., KÄRGEL, E. & SCHUNCK, W.-H. 1996. Characterization of then-Alkane and Fatty Acid Hydroxylating Cytochrome P450 Forms 52A3 and 52A4. *Archives of biochemistry and biophysics,* 328**,** 245-254.

SCHEPS, D., MALCA, S. H., HOFFMANN, H., NESTL, B. M. & HAUER, B. 2011. Regioselective ω-hydroxylation of medium-chain n-alkanes and primary alcohols by CYP153 enzymes from Mycobacterium marinum and Polaromonas sp. strain JS666. *Organic & Biomolecular Chemistry,* 9**,** 6727-6733.

SCHEWE, H., HOLTMANN, D. & SCHRADER, J. 2009. P450 BM-3-catalyzed whole-cell biotransformation of α-pinene with recombinant Escherichia coli in an aqueous–organic two-phase system. *Applied microbiology and biotechnology,* 83**,** 849-857.

SCHMID, A., SONNLEITNER, B. & WITHOLT, B. 1998. Medium chain length alkane solvent-cell transfer rates in two-liquid phase, Pseudomonas oleovorans cultures. *Biotechnology and bioengineering,* 60**,** 10-23.

SCHMITZ, A. 2012. *General Chemistry: Principles, Patterns and Applications* [Online]. Saylor Academy. Available: https://saylordotorg.github.io/text_general-chemistry-principles-patterns-and-applications-v1.0/index.html [Accessed 27 June 2020].

SCHREWE, M., JULSING, M. K., BUEHLER, B. & SCHMID, A. 2013. Whole-cell biocatalysis for selective and productive C–O functional group introduction and modification. *Chemical Society Reviews,* 42**,** 6346-6377.

SCHREWE, M., JULSING, M. K., LANGE, K., CZARNOTTA, E., SCHMID, A. & BÜHLER, B. 2014. Reaction and catalyst engineering to exploit kinetically controlled whole-cell multistep biocatalysis for terminal FAME oxyfunctionalization. *Biotechnology and bioengineering,* 111**,** 1820-1830.

SCHWERT, G. W. 1969. Use of integrated rate equations in estimating the kinetic constants of enzyme-catalyzed reactions. *Journal of Biological Chemistry,* 244**,** 1278-1284.

SCILLIPOTI, J., NIOI, C., MARTY, A., CAMY, S. & CONDORET, J.-S. 2017. Prediction of conversion at equilibrium for lipase esterification in two-phase systems. *Biochemical Engineering Journal,* 117**,** 162-171.

SEIFERT, A. & PLEISS, J. 2009. Identification of selectivity-determining residues in cytochrome P450 monooxygenases: A systematic analysis of the substrate recognition site 5. *Proteins: Structure, Function, and Bioinformatics,* 74**,** 1028-1035.

SHANKLIN, J., ACHIM, C., SCHMIDT, H., FOX, B. G. & MÜNCK, E. 1997. Mössbauer studies of alkane ω-hydroxylase: evidence for a diiron cluster in an integral-membrane enzyme. *Proceedings of the National Academy of Sciences,* 94**,** 2981-2986.

SHILOV, A. E. & SHUL'PIN, G. B. 1997. Activation of C− H bonds by metal complexes. *Chemical Reviews,* 97**,** 2879-2932.

SMITH, J. M. 1950. Introduction to chemical engineering thermodynamics. ACS Publications.

SPEIGHT, J. G. 2018. *Reaction Mechanisms in Environmental Engineering: Analysis and Prediction*, Butterworth-Heinemann.

SRINIVASAN, B. 2022. A guide to the Michaelis–Menten equation: steady state and beyond. *The FEBS journal,* 289**,** 6086-6098.

STAIJEN, I. E., VAN BEILEN, J. B. & WITHOLT, B. 2000. Expression, stability and performance of the three-component alkane mono-oxygenase of Pseudomonas oleovorans in Escherichia coli. *European Journal of Biochemistry,* 267**,** 1957-1965.

STRAATHOF, A., RAKELS, J. & HEIJNEN, J. 1992. Enzyme kinetics in monophasic and biphasic aqueous-organic systems. *Progress in Biotechnology.* Elsevier.

STRAATHOF, A. J. 2003. Enzymatic catalysis via liquid–liquid interfaces. *Biotechnology and Bioengineering,* 83**,** 371-375.

SWAINSTON, N., BAICI, A., BAKKER, B. M., CORNISH-BOWDEN, A., FITZPATRICK, P. F., HALLING, P., LEYH, T. S., O'DONOVAN, C., RAUSHEL, F. M. & RESCHEL, U. 2018.

STRENDA DB: enabling the validation and sharing of enzyme kinetics data. *The FEBS Journal,* 285**,** 2193-2204.

TÁNCSICS, A., BENEDEK, T., SZOBOSZLAY, S., VERES, P. G., FARKAS, M., MÁTHÉ, I., MÁRIALIGETI, K., KUKOLYA, J., LÁNYI, S. & KRISZT, B. 2015. The detection and phylogenetic analysis of the alkane 1-monooxygenase gene of members of the genus Rhodococcus. *Systematic and applied microbiology,* 38**,** 1-7.

THOMAS, P. E., RYAN, D. & LEVIN, W. 1976. An improved staining procedure for the detection of the peroxidase activity of cytochrome P-450 on sodium dodecyl sulfate polyacrylamide gels. *Analytical biochemistry,* 75**,** 168-176.

TIPTON, K. F., ARMSTRONG, R. N., BAKKER, B. M., BAIROCH, A., CORNISH-BOWDEN, A., HALLING, P. J., HOFMEYR, J.-H., LEYH, T. S., KETTNER, C. & RAUSHEL, F. M. 2014. Standards for Reporting Enzyme Data: The STRENDA Consortium: What it aims to do and why it should be helpful. *Perspectives in Science,* 1**,** 131-137.

TSUKAGOSHI, N. & AONO, R. 2000. Entry into and release of solvents by Escherichia coli in an organic-aqueous two-liquid-phase system and substrate specificity of the AcrAB-TolC solvent-extruding pump. *Journal of Bacteriology,* 182**,** 4803-4810.

TUFVESSON, P., JENSEN, J. S., KROUTIL, W. & WOODLEY, J. M. 2012. Experimental determination of thermodynamic equilibrium in biocatalytic transamination. *Biotechnology and Bioengineering,* 109**,** 2159-2162.

URLACHER, V. B. & GIRHARD, M. 2012. Cytochrome P450 monooxygenases: an update on perspectives for synthetic application. *Trends in biotechnology,* 30**,** 26-36.

URLACHER, V. B., LUTZ-WAHL, S. & SCHMID, R. D. 2004. Microbial P450 enzymes in biotechnology. *Applied microbiology and biotechnology,* 64**,** 317-325.

VAN BEILEN, J. B., DUETZ, W. A., SCHMID, A. & WITHOLT, B. 2003a. Practical issues in the application of oxygenases. *Trends in biotechnology,* 21**,** 170-177.

VAN BEILEN, J. B. & FUNHOFF, E. G. 2005. Expanding the alkane oxygenase toolbox: new enzymes and applications. *Current Opinion in Biotechnology,* 16**,** 308-314.

VAN BEILEN, J. B. & FUNHOFF, E. G. 2007. Alkane hydroxylases involved in microbial alkane degradation. *Applied Microbiology and Biotechnology,* 74**,** 13-21.

VAN BEILEN, J. B., FUNHOFF, E. G., VAN LOON, A., JUST, A., KAYSSER, L., BOUZA, M., HOLTACKERS, R., RÖTHLISBERGER, M., LI, Z. & WITHOLT, B. 2006. Cytochrome P450 alkane hydroxylases of the CYP153 family are common in alkane-degrading eubacteria lacking integral membrane alkane hydroxylases. *Applied and Environmental Microbiology,* 72**,** 59-65.

VAN BEILEN, J. B., LI, Z., DUETZ, W. A., SMITS, T. H. & WITHOLT, B. 2003b. Diversity of alkane hydroxylase systems in the environment. *Oil & gas science and technology,* 58**,** 427-440.

VAN TOL, J. B. A., JONGEJAN, J. A., DUINE, J. A., KIERKELS, H. G., GELADÉ, E. F., MOSTERD, F., VAN DER TWEEL, W. J. & KAMPHUIS, J. 1995. Thermodynamic and kinetic parameters of lipase-catalyzed ester hydrolysis in biphasic systems with varying organic solvents. *Biotechnology and bioengineering,* 48**,** 179-189.

VOGES, M., ABU, R., GUNDERSEN, M. T., HELD, C., WOODLEY, J. M. & SADOWSKI, G. 2017. Reaction equilibrium of the ω-transamination of (S)-phenylethylamine: Experiments and ePC-SAFT modeling. *Organic Process Research & Development,* 21**,** 976-986.

WACHTMEISTER, J. & ROTHER, D. 2016. Recent advances in whole cell biocatalysis techniques bridging from investigative to industrial scale. *Current opinion in biotechnology,* 42**,** 169-177.

WALTER, H. 2012. *Partitioning in Aqueous Two–Phase System: Theory, Methods, Uses, and Applications to Biotechnology*, Elsevier.

WANGLER, A., BUNSE, M. J., SADOWSKI, G. & HELD, C. 2018. Thermodynamic activity-based Michaelis constants. *Kinetics of Enzymatic Synthesis*.

WESCOTT, C. R. & KLIBANOV, A. M. 1993. Predicting the solvent dependence of enzymic substrate specificity using semiempirical thermodynamic calculations. *Journal of the American Chemical Society,* 115**,** 10362-10363.

WHITE, B. E., FENNER, C. J., SMIT, M. S. & HARRISON, S. T. 2017. Effect of cell permeability and dehydrogenase expression on octane activation by CYP153A6-based whole cell Escherichia coli catalysts. *Microbial cell factories,* 16**,** 156.

WHITEHOUSE, C. J., BELL, S. G. & WONG, L.-L. 2012. P450 BM3 (CYP102A1): connecting the dots. *Chemical Society Reviews,* 41**,** 1218-1260.

WIENER, M. C. & HORANYI, P. S. 2011. How hydrophobic molecules traverse the outer membranes of Gram-negative bacteria. *Proceedings of the National Academy of Sciences,* 108**,** 10929-10930.

WIENKE, G. & GMEHLING, J. 1998. Prediction of octanol-water partition coefficients, Henry coefficients and water solubilities using UNIFAC. *Toxicological & Environmental Chemistry,* 65**,** 57-86.

WILTSCHI, B., CERNAVA, T., DENNIG, A., CASAS, M. G., GEIER, M., GRUBER, S., HABERBAUER, M., HEIDINGER, P., ACERO, E. H. & KRATZER, R. 2020. Enzymes revolutionize the bioproduction of value-added compounds: From enzyme discovery to special applications. *Biotechnology Advances***,** 107520.

YADAV, G. D. & MAGADUM, D. B. 2017. Kinetic Modelling of Enzyme Catalyzed Biotransformation Involving Activations and Inhibitions. *Enzyme Inhibitors and Activators***,** 73.

YANG, H., LI, J., SHIN, H.-D., DU, G., LIU, L. & CHEN, J. 2014. Molecular engineering of industrial enzymes: recent advances and future prospects. *Applied microbiology and biotechnology,* 98**,** 23-29.

YANG, K., CAI, Z., JAISWAL, A., TYAGI, M., MOORE, J. S. & ZHANG, Y. 2016. Dynamic Odd–Even Effect in Liquid n-Alkanes near Their Melting Points. *Angewandte Chemie International Edition,* 55**,** 14090-14095.

7 Appendix for experimental work and results

7.1 Heat Shock Transformation of BL21 E. coli with the CYP153A13 gene and DNA agarose gel electrophoresis

This was used for microorganism preparation (Chapter 3.1.1). *E.coli* BL21(DE3) competent cells were prepared by monitoring the growth of a culture at 37˚C, 200 rpm until an OD_{600} measurement of 0.4 to 0.6. At this point, cells were placed on ice at 2-4˚C for 30 minutes. Cells were centrifuged at 3000 rpm for five minutes at 4˚C. The cell pellet was gently and slowly resuspended in 6 mL of filter sterilised 50 mM $CaCl_2$ solution supplemented with 20% glycerol. Resultant aliquots of competent cells were stored at -60˚C.

A volume of 1 µL of 10 ng/µL stock of CYP153A13pET-28b(+) plasmid was added to 300 µL of defrosted competent BL21 *E. coli* cells and left on ice for 30 minutes. The sample tube containing the cell and DNA mixture was incubated at 37˚C for 90 seconds and then placed on ice for two minutes. A volume of 700 µL LB was added to the sample and incubated at 37˚C for 30 minutes.

A volume of 100 µL of cell suspension and plasmid DNA mixture sample was placed on solid selection media agar plates containing 30 µg/mL kanamycin. The remaining volume of sample was centrifuged for five minutes and 800 µL supernatant was discarded. The remaining cell pellet was gently resuspended in the remaining 100 µL volume of media and pipetted onto solid media selection plates containing 30 µg/mL kanamycin. Selection plates containing samples were incubated overnight at room temperature. Controls were included in the experiment, such as the transformant control to ensure that competent cells were functional, the cell viability control, antibiotic resistance control which ensure that strains did not have previous antibiotic resistance. Positive recombinant transformant single colonies were picked and grown overnight in a volume of 10 mL LB media containing 30 µg/mL kanamycin. Glycerol stocks of positive transformant cells were stored at 60˚C.

Once the heat shock transformation was complete, the PureYield™ Plasmid Miniprep System was used to extract the inserted plasmid DNA. The centrifugation protocol was used. Recombinant cell pellet was resuspended in a volume of 600 µL dH_2O. A volume of 100 µL of Cell Lysis Buffer was added and mixed for a maximum of two minutes, until the mixture changed from opaque to blue, which indicated complete cell lysis. A volume of 350 µL of Neutralisation Solution with a temperature of 4 to 8 ˚C was added and the tube inverted until the mixture turned yellow, and a precipitate was formed.

The sample was centrifuged at maximum speed in a microcentrifuge for three minutes. The supernatant was transferred to a PureYield™ minicolumn. The minicolumn was placed into a PureYield™ collection tube and centrifuged for 15 seconds at maximum speed. The flow through was discarded and the minicolumn was placed into the same collection tube. A volume of 200 µL of Endotoxin Removal Wash was added to the minicolumn and centrifuged at maximum speed for 15 seconds. A volume of 400 µL of Column Wash Solution was added to the minicolumn and centrifuged at maximum speed for 30 seconds.

The minicolumn was transferred to a clean 1.5 mL microcentrifuge and 30 µL of elution buffer was added directly to the minicolumn matrix and allowed to stand for one minute at room temperature. The sample was centrifuged at maximum speed for 15 seconds to elute the plasmid DNA. Samples of the plasmid DNA were stored.

The isolated plasmid DNA underwent diagnostic restriction enzyme digest to validate the presence of the CYP153A13Fusion insert and vector in the BL21 *E. coli* host cells. The restriction enzyme cleavage sites present on the plasmid were identified to be HindIII and NdeI. A volume of 17 µL plasmid DNA, 1 µL HindIII enzyme and 2 µL buffer were combined into a sterile Eppendorf tube. The sample was mixed

by gentle inversion for five minutes. An additional 2.5 µL buffer, 0.5 µL NdeI and 2 µL dH_2O were added and allowed to incubate for one hour at 37˚C.

Samples were analysed using agarose gel electrophoresis. A mass of 0.56 g agarose powder was added to 84 mL 1×TAE Buffer and boiled in a microwave. A volume of 1 µL Ethidium Bromide was added when the mixture cooled down to approximately 50˚C. This mixture was poured into a gel tray with the well comb in place and allowed to solidify at room temperature.

A volume of 5 µL KAPA Express Ladder was added to the first well (Well A). A volume of 20 µL undigested plasmid DNA and 4 µL loading dye was added to the second well (Well B). A volume of 4 µL loading dye and 25 µL digested plasmid DNA were added to the third, fourth and fifth wells (Well C, D and E). TAE running buffer was added to cover the surface of the gel and 90 V was applied to run the gel. The gel was visualised under UV light (Figure 5).

7.2 Standard curve of dry cell weight versus optical density at 600 nm

This method enabled easier preparation of the bioreaction mixture suspension because dry cell weight takes 48 hours to measure, whereas OD_{600} takes a few minutes. Therefore, the following standard curve was prepared (Chapter 3.1.6). The preparation of the BRM required the suspension of a specific cell concentration to obtain a certain mass of dry cell weight (DCW) per volumetric unit of the sodium phosphate buffer. Cell turbidity can be measured quickly (as explained in the previous paragraph), while DCW takes at least 48 hours. Therefore, a standard curve was constructed to establish the relationship between DCW and OD_{600}, showing an error of 5 and 4%, respectively (Appendix 7.2). The OD_{600} of samples was measured quickly (c_1) to determine the correlated DCW by using the standard curve. The final desired concentration (c_2) for BRM in DCW measured in $g_{DCW} \cdot L_{BRM}^{-1}$ was prepared according to the initial cell concentration of the recombinant (c_1)

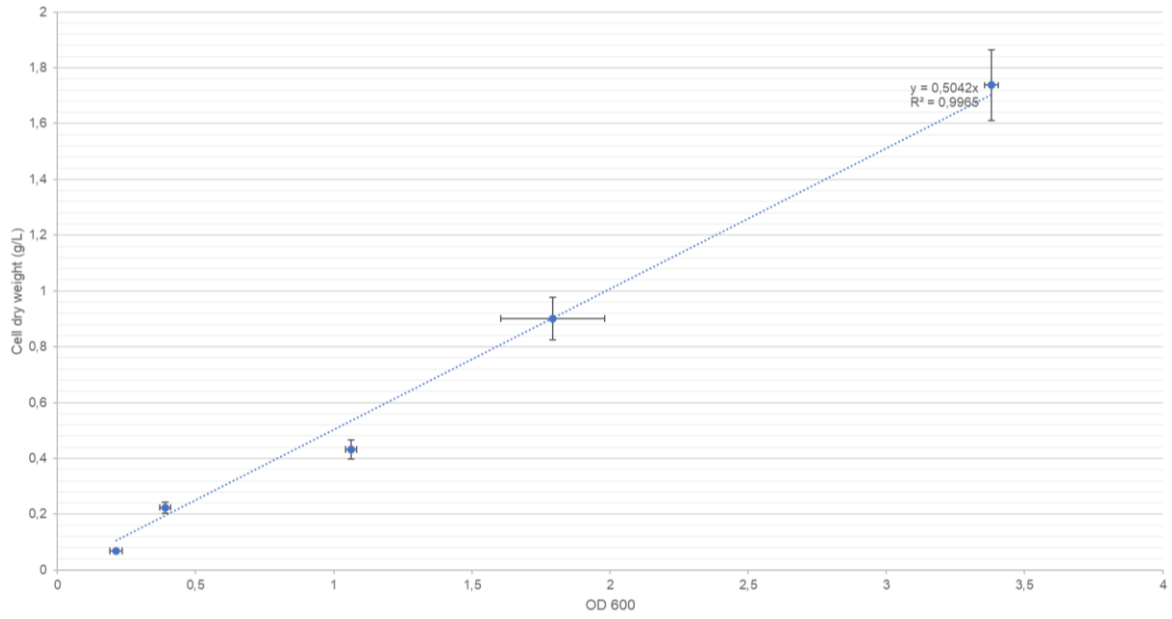


Figure 7.1 Standard curve of dry cell weight versus optical density measured at 600 nm (OD600). Error bars show average error in duplicate analysis.

7.3 Table of CYP concentrations in biological repeats measured using CO-difference spectra

CO-difference spectral analyses measured the CYP concentrations in the BRM. CYP concentrations were taken as nominal average ± SE 95% CI of three other biological repeats with a relative standard error of 10% to 15%. This allowed the calculation of turnover frequency etc (See Chapter 7.1.6).

Table 7.1 CYP concentrations measured from the CO-difference spectra of different biological repeats

	CYP concentration ($\mu mol_{P450} \cdot g_{DCW}^{-1}$)	
Biological Repeat (BR)	CYP153A6	CYP153A13
BR1	0.0775	0.0392
BR2	0.0765	0.0324
BR3	0.1037	0.0543
Average ± SE 95% CI	0.0859 ± 0.0089	0.0420 ± 0.0065
Relative Standard Error (%)	10	15

7.4 Gas chromatography specifications

Gas chromatography was used to measure primary alcohol concentration using the following specifications (Chapter 3.2.3)

Table 7.2 Gas chromatography specifications

Gas chromatograph	Varian 3900 series
Detector	Flame ionization detector (FID)
Column	Non-polar Varian factorFour™ (VF-1ms)
Internal Column material	Dimethylpolysiloxane
Column Dimensions	15 m × 0.25 mm
Inlet temperature	280˚C
Column temperature	120˚C for 5 minutes, ramped up to 280˚C in increments of 20˚C per min followed by a seven-minute isotherm
Carrier gas	15 $mL \cdot min^{-1}$ Nitrogen

7.5 Standard curves of primary alcohol products

For primary alcohol concentration measurement, standard curves were used (Chapter 3.2.3). Primary alcohol analysed generated peaks at specific retention times in the gas chromatograms. Subsequently, primary alcohol concentrations were calculated using standard curves ranging from primary alcohol concentrations of 20mM to 100mM. The solvent used to dilute the primary alcohol contained an identical concentration (2mM) of internal standard as the solvent used for sacrificial extractions of the BRM. For analysis of 1-octanol and 1-nonanol, and 1-decanol internal standard was used and for 1-decanol analysis a 1-octanol internal standard was used. The y-axis of the standard curve comprised the ratio of the peak area of the primary alcohol product to the peak area of the chosen internal standard, and the x-axis comprised the independent variable which was alcohol concentration. The standard curves constructed for the different alcohol products had a 3% and less average relative standard error of the mean.

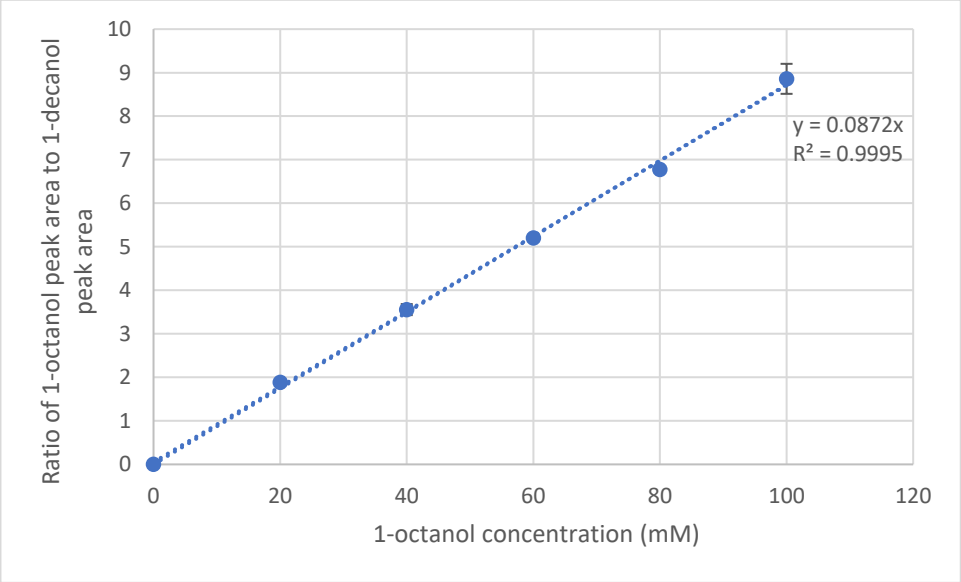


Figure 7.2 1-octanol standard curve constructed using defined 1-octanol concentrations dissolved in ethyl acetate containing 2 mM 1-decanol internal standard. The error bars indicate the standard error of mean of three repeats.

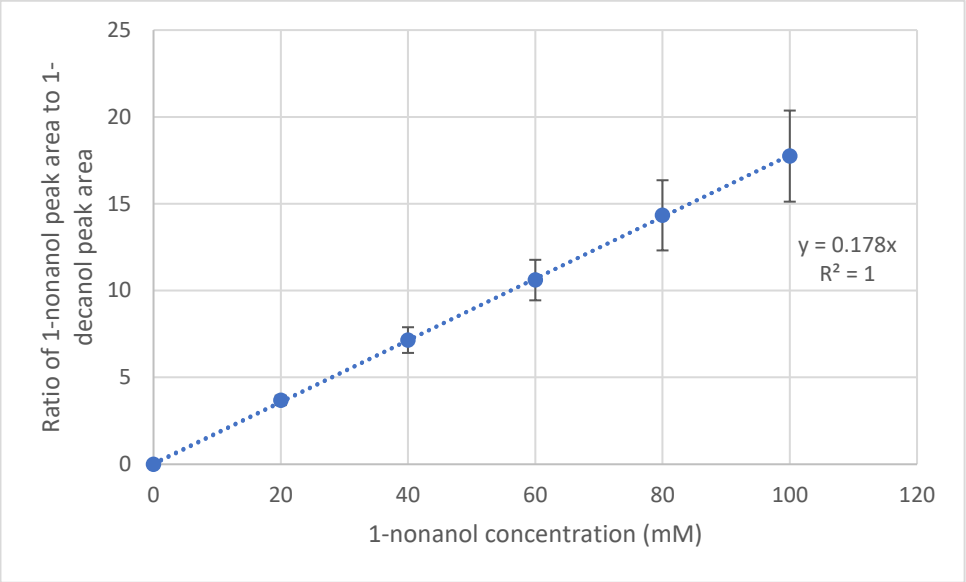


Figure 7.3 1-nonanol standard curve constructed using defined 1-nonanol concentrations dissolved in ethyl acetate containing 2 mM 1-decanol internal standard. The error bars indicate the standard error of mean of two repeats.

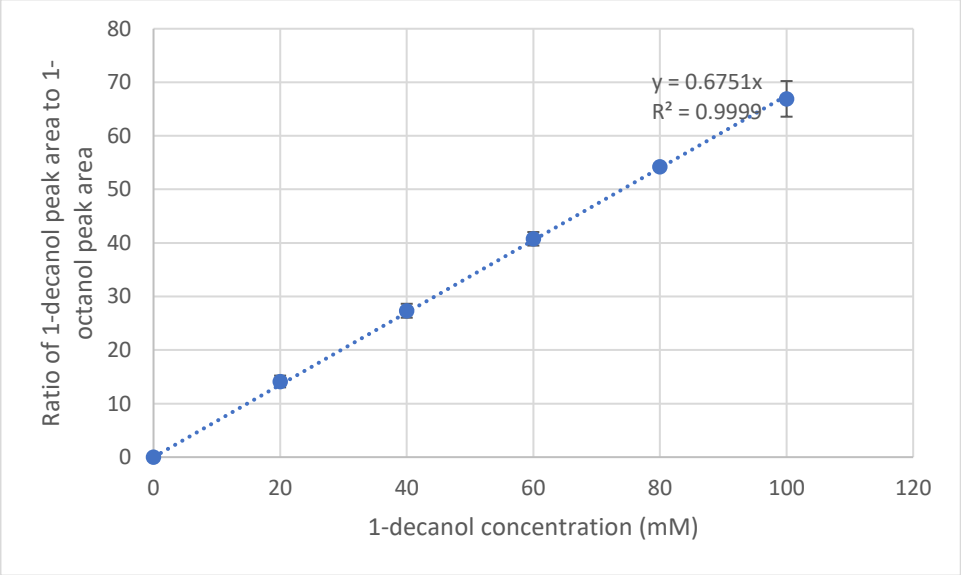
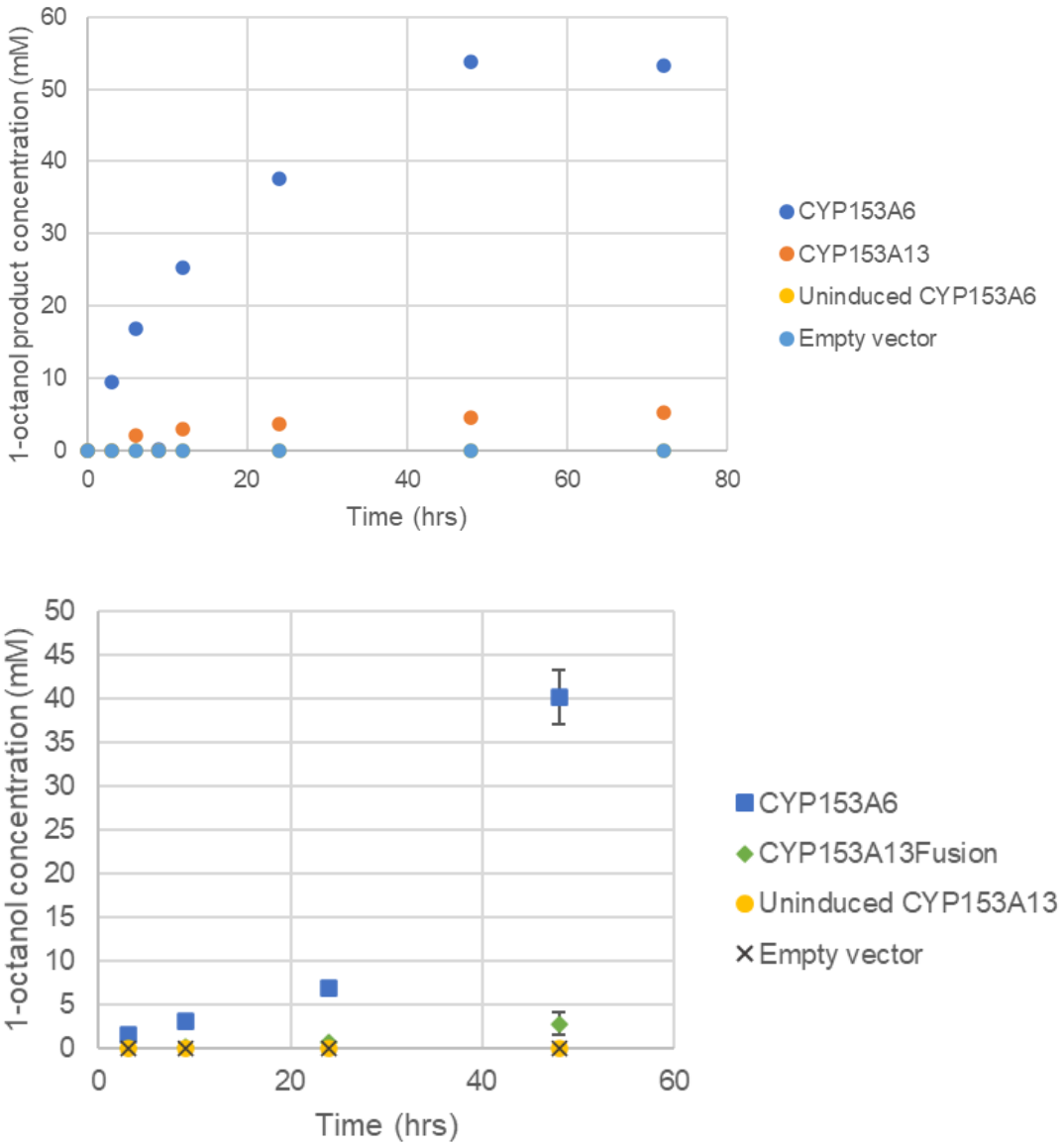


Figure 7.4 1-decanol standard curve constructed using defined 1-nonanol concentrations dissolved in ethyl acetate containing 2 mM 1-octanol internal standard. The error bars indicate the standard error of the mean of three repeats.

7.6 Results showing negative controls (Uninduced CYP153A6, uninduced CYP153A13 and empty vector)



7.7 Data processing and metrics (c*Change standards)

The raw data was uploaded to ZivaHub and can be found in the link: https://doi.org/10.25375/uct.23733378.v1. The following variables were calculated using the raw data, following the c*Change Centre of Excellence in Catalysis Research standard of reporting.

Glossary – SPECIFIC ACTIVITY / TOF

$$\text{TOF} = \frac{\text{moles product}}{\text{moles active sites}} \cdot \frac{1}{\text{time}}$$

Assume: moles product = moles of substrate consumed

	PRODUCT	ACTIVE SITE	units
BIO	1-octanol	P450	$[mol_{1\text{-octanol}}/mol_{P450}/h]$
CHEM	1-octanol & oxygenates	Me ion	$[mol_{1\text{-octanol}}/mol_{ions}/h]$

Glossary – PRODUCTIVITY / VOLUMETRIC RATE

$$\text{rate} = \frac{\text{moles product}}{\text{working volume}} \cdot \frac{1}{\text{time}}$$

Based on the working volume (WV) available for reaction and NOT reactor size nor only aqueous volume i.e. it does not exclude the solvent

	PRODUCT	units
BIO	oxygenates	$[mol/L_{WV}/h]$
CHEM	oxygenates	$[mol/L_{WV}/h]$

Glossary – (BIO)CATALYTIC EFFICIENCY

$$\text{efficiency} = \frac{\text{mass product}}{\text{mass catalyst}}$$

where for the bio-system, mass catalyst = DCW

	PRODUCT	units
BIO	1-octanol	$[g_{1\text{-octanol}}/g_{DCW}]$
CHEM	1-octanol	$[g_{1\text{-octanol}}/g_{catalyst}]$

Glossary – Max. Specific Biocatalytic Activity

- Similar to TOF but per amount of whole cell rather than P450

$$\text{bioctalytic activity} = \frac{\mu\text{mol product}}{\text{mass catalyst}} \cdot \frac{1}{\text{time}}$$

	PRODUCT	units
BIO	1-octanol	$[\mu mol_{1\text{-octanol}}/g_{DCW}/min]$
CHEM	oxygenates	$[\mu mol_{oxygenates}/g_{catalyst}/min]$

Glossary – Specific Intracellular P450 Concentration

$$\text{P450 concentration} = \frac{\mu\text{mol P450}}{\text{dry cell weight}}$$

	PRODUCT	units
BIO	P450	$[\mu mol_{P450}/g_{DCW}]$
CHEM		

7.8 Gas chromatograms for activity tests of CYP153A13 fusion extracts

These gas chromatograms were obtained for purified CYP153A13 hydroxylation tests for 24 hours reaction time (Chapter 4.1.1). The top figure is cleared lysate, middle is desalted and bottom figure is dialysed CYP153A13, respectively.

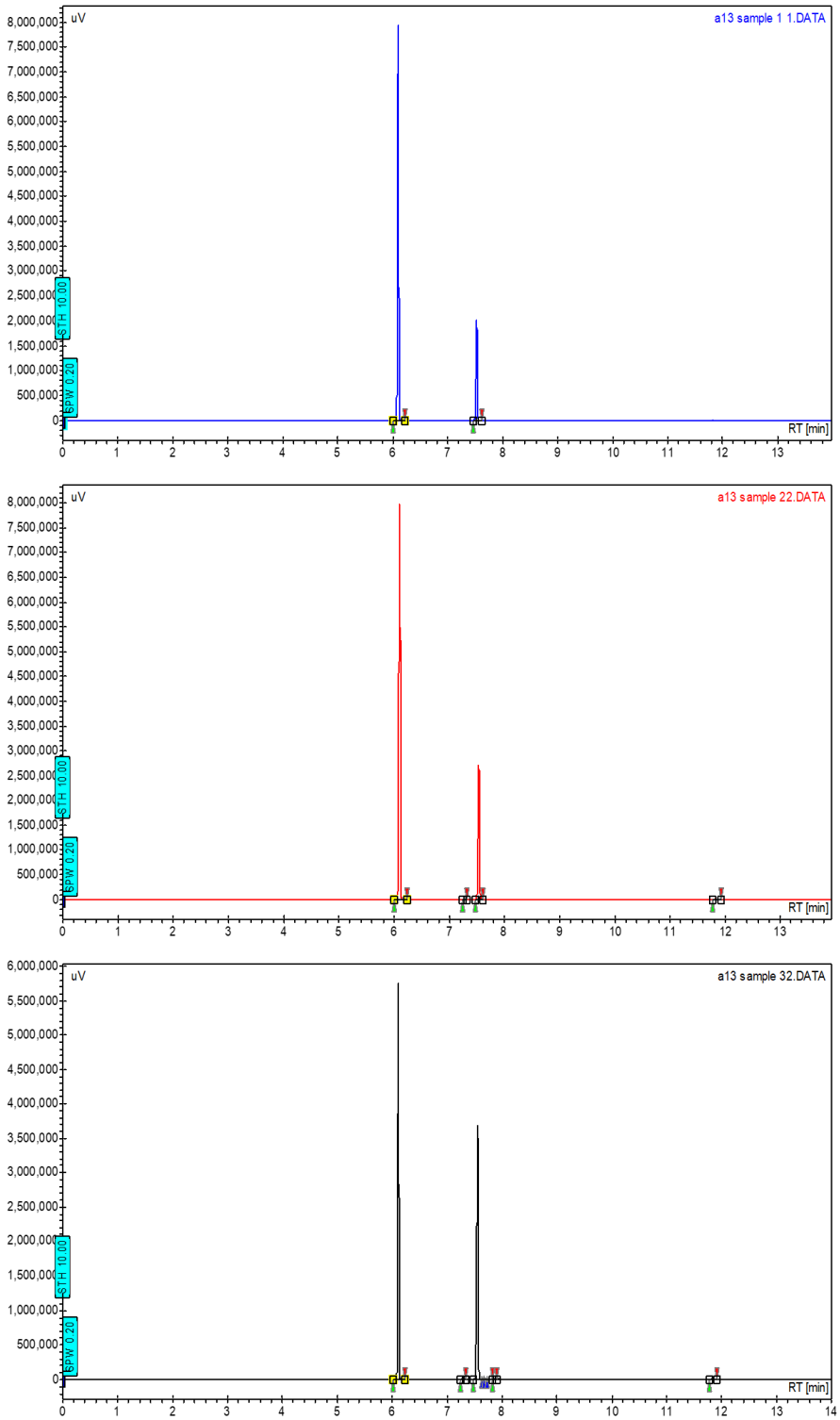


Figure 7.5 Gas chromatograms of cleared lysate (top) desalted (middle) and dialysed (bottom) CYP153A13 Fusion extracts illustrating that 1-octanol was not present after 24 hours of exposure to octane. Peaks of 1-octanol concentration should be observed at 11.9 min.

7.9 Gas chromatograms for activity tests of CYP102A1

These gas chromatograms were obtained for purified CYP102A1 hydroxylation tests (Chapter 4.1.1). The top is cleared lysate, middle is desalted, and bottom is dialysed. CYP102A1 extracts showing 2-octanol after 24 hours of exposure to octane.

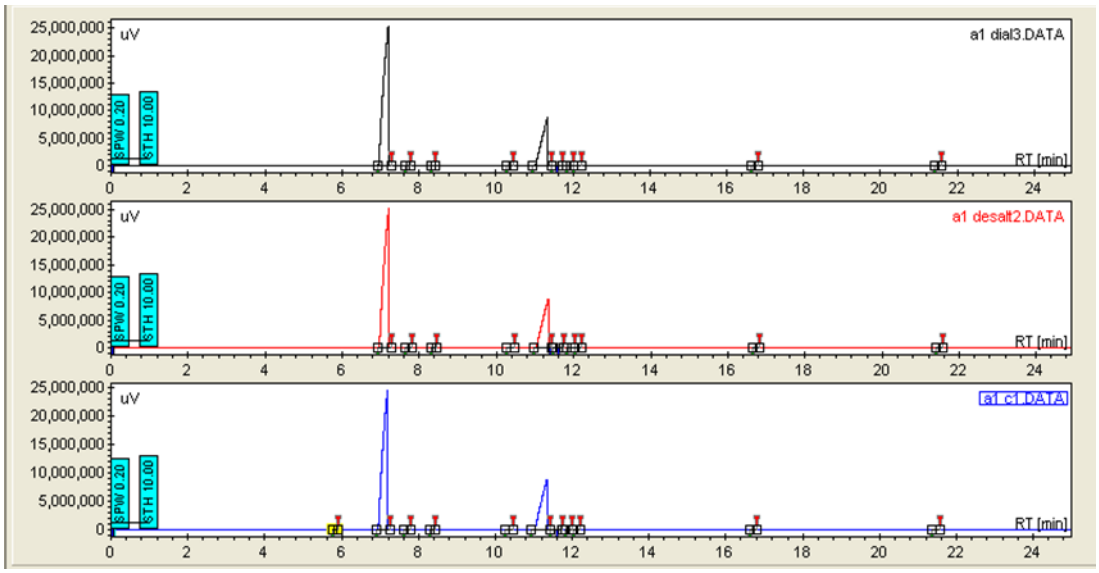


Figure 7.6 Gas chromatograms of crude lysate (top) desalted (middle) and dialysed (bottom) CYP102A1 extracts illustrating that 2-octanol was present after 24 hours of exposure to octane. 2-octanol is seen at the 17 min peak.

7.10 Product concentration measured using the GC for high cell and low cell density cultures

The tables below show the experimental 1-octanol product concentration measured (raw data) that analysed in Chapter 4.1.2 for high and low cell density biohydroxylation experimentation. These were measured using GC (Chapter 3.2.3)

Table 7.3 Product concentration (μM) measured using the GC for high cell and low cell density cultures.

	CYP153A6 LD	CYP153A6 HD	CYP153A13 LD	CYP153A13 HD
Time (hrs)	Repeat 1	Repeat 1	Repeat 1	Repeat 1
24	15799		5553	19105
48	57859	627949	7349	24651