Directed Evolution of P450cin for Mediated Electron Transfer

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften der RWTH Aachen University zur Erlangung des akademischen Grades eines Doktorin der Naturwissenschaften genehmigte Dissertation

vorgelegt von

Ketaki Deepak Belsare, M.Tech.

aus Nasik, Indien

Berichter: Universitätsprofessor Dr. Ulrich Schwaneberg

Universitätsprofessor Dr. Lothar Elling

Tag der mündlichen Prüfung: 29. Oktober 2014

Diese Dissertation ist auf den Internetseiten der Hochschulbibliothek online verfügbar

To my parents and sister for their love,
endless support and encouragement

ACKNOWLEDGEMENTS

ACKNOWLEDGEMENTS

PUBLICATIONS

Parts of this thesis have been published

Publication:

Ketaki D. Belsare, Anna Joëlle Ruff, Ronny Martinez, Amol V. Shivange, Hemanshu Mundhada, Dirk Holtmann, Jens Schrader and Ulrich Schwaneberg, P-LinK: A Method for Generating Multicomponent Cytochrome P450 Fusions with Variable Linker Length, *BioTechniques* 57: 13-20; July 2014; doi 10.2144/000114187.

Poster contributions:

Ketaki D. Belsare, Amol V. Shivange, Hemanshu Mundhada, Dirk Holtmann, Jens Schrader, and Ulrich Schwaneberg; Directed Evolution of the Cytochrome P450cin for Mediated Electron Transfer; Biotrans, Sicily, Italy 2011.

Ketaki D. Belsare, Anna Joelle Ruff , Ronny Martinez , Dirk Holtmann, Jens Schrader and Ulrich Schwaneberg; A Versatile and Rapid Method for Generation of Fusion Proteins; International Conference on Cytochrome P450- Biochemistry, Biophysics and Structure; Seattle, Washington, USA 2013.

Part of this thesis will be published

Ketaki D. Belsare , Anna Joelle Ruff , Ronny Martinez , Anders Magnusson, Dirk Holtmann, Jens Schrader and Ulrich Schwaneberg; Directed Evolution of P450cin Fusion Protein for Mediated Electron Transfer (manuscript in preparation).

Ketaki D.Belsare, Anna Joelle Ruff, Ronny Martinez, and Ulrich Schwaneberg; Insight in the Role of Linker Region in P450cin Fusion Protein (manuscript in preparation).

ABSTRACT

Cytochrome P450s represent a promising class of biocatalysts which has emerged as an alternative and eco-friendly synthesis route for fine chemicals and pharmaceuticals in the past decade. Protein engineering of cytochrome P450s offer opportunity to generate tailor-made biocatalysts that match industrial demands. The work described in this thesis tackles two major hurdles while employing P450s in cell free reaction systems – multi-component nature of P450s and the necessity of NAD(P)H, an expensive co-factor for the conversion of substrate to product. The work described in this thesis has been performed using cytochrome P450cin.

Cytochrome P450cin is a multi-component bacterial monooxygenase consisting of three domains namely CinA (heme), CinB (cindoxin reductase, FAD) and CinC (cindoxin, FMN). It stereo specifically converts 1,8-cineole to 2-β-hydroxy-1,8-cineole using NADPH as the reduction equivalent. CinB is usually replaced in activity measurements by the *E. coli* NADPH-ferredoxin (flavodoxin) reductase (Fpr) since CinB is not expressible in an active form.

The first part of the thesis describes a novel, technically simple and enzyme free method (P-LinK-Protein fusion with variable linker insertion) for generation of fusion proteins with variable linker length. Using P-Link CinC was fused to the C terminal end of CinA through sixteen different linker lengths in a single experiment employing three PCRs. Enzymatic activity screening of the linker library revealed that a linker length between 8 and 12 amino acids is optimal in terms of 1,8-cineole hydroxylation activity. The highest activity was observed for fusion protein with a linker length of 10 amino acids.

The second part of the thesis describes characterization of the CinA-CinC fusion protein with a linker length of 10 amino acids. Investigation using computational and experimental approach indicates that the electron transfer pathway in the P450cin fusion protein is predominantly intermolecular and that the length of the linker separating the fused domains is crucial in determining the activity of the fusion protein.

In the third part of the thesis, the CinA-linker-10 aa-CinC fusion protein was engineered using random and rational protein engineering for improvement in 1,8-cineole hydroxylation activity using alternative co-factor system. The main goal of the protein engineering campaign was to replace NADPH in cell free reactions with alternative co-factor

system. Two different alternative co-factor systems have been investigated (Zn/Co(III)sepulchrate and the second was a bioelectrocatalytic approach employing platinum electrodes and Co(III)sepulchrate). For the approach using Zn/Co(III)sepulchrate a mutant KB8 carrying four mutations (Q385H/V386S/T77N/L88R) showed a 3.8 times improved catalytic efficiency (k_{cat} /K_m) (k_{eff} = 27.77 µM min^{-1} mM^{-1}) compared to the starting wild type variant (CinA-linker10aa-CinC) (k_{eff} = 7.4 µM min^{-1} mM^{-1}). k_{cat} value for mutant KB8 (2.5 µM min^{-1}) improved by about 4-fold as compared to the wild type (0.67 µM min^{-1}) with no significant change in the K_m value (90 µM). Total turnover numbers for the mutant KB8 were determined to be 1170 and 480 using 5 mM and 1 mM 1,8-cineole concentration respectively. A rational protein engineering approach was performed and amino acid residues 79, 80, and 81 were selected based on docking studies of Co(III)sepulchrate in the active site of the P450cin fusion protein with 10 amino acid linker length. The sites 79, 80, and 81 completed a putative path of electron transfer from the Co(III)sepulchrate to the active site of the enzyme. A mutant KB9 (Q385H/V386S/P79G/R80D/Y81S) with 1.5 fold improvement in activity was identified after screening of simultaneous site saturation library at the selected positions (amino acid residues 79, 80, and 81).

Preliminary investigation using electrobiocatalytic reactions employing platinum and Co(III)sepulchrate displayed highest product formation (61 µM µM^{-1} P450) in a reaction time of 6 h for the mutant KB8 which is approximately 3 times higher than that obtained for WT variant (CinA-linker10aa-CinC) (23 µM µM^{-1} P450).

ABBREVIATIONS

ALA	5-Aminolevulinic Acid Hydrochloride
Co(III)sepulchrate	Cobalt (III) sepulchrate
CYP	Cytochrome P450 Monooxygenase
dATP	Deoxyadenosine triphosphate
dGTP	Deoxyguanosine triphosphate
ddH$_2$O	Double distilled water
DEAe	Diethylethanolamine
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
E. coli	*Escherichia coli*
e.g.	*exempli gratia* (for example)
FAD	Flavin Adenine Dinucleotide
FID	Flame ionization detector
FMN	Flavin Mononucleotide
GC	Gas Chromatography
h	Hour(s)
IPTG	Isopropyl-β-D-thiogalactopyranoside
Kan	Kanamycin
kb	Kilobase(s)
kDa	Kilo Dalton
LB	Luria-Bertani Medium
min	Minutes
MTP	Microtiter Plate
NADPH	Nicotinamide adenine dinucleotide phosphate
n.d.	Not determined
OD$_{600}$	Optical density of viable bacteria at 600 nm
PCR	Polymerase Chain Reaction
rpm	Revolutions per minute
SDM	Site Directed Mutagenesis
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SeSaM	Sequence Saturation Mutagenesis
SOC	Super optimal broth with catabolic repressor
TB	Terrific Broth Medium
TLC	Thin layer chromatography
UV	Ultra Violet
WT	Wild type

CONTENTS

CONTENTS

CONTENTS

CONTENTS

CONTENTS

1. INTRODUCTION

1.1 Enzymes: Towards "Green Chemistry"

Enzymes are highly specialized proteins, present in every living organism and catalyze innumerable different chemical reactions necessary to sustain life. In the field of chemistry, enzymes are capable of performing chemical reactions with high *regio* and *stereo* selectivity in comparison to chemical catalysts, act under mild conditions of temperature and pressure as opposed to chemical catalysts and generate practically no toxic waste products (Cooper, 2000). Thus they provide an environmentally friendly "green chemistry" approach to perform preparative chemical synthetic reactions. United States Environmental Protection Agency defines green chemistry as the design of chemical products and processes that reduce or eliminate the use or generation of hazardous substances. It applies to the use or generation of a chemical product including its design, manufacture, use and disposal. The impressive repertoire of these biocatalysts can be utilized as an alternative to chemical synthesis (Koeller and Wong, 2001).

However, enzymes are extremely well suited for supporting life but are usually not very suitable for applications in industry and technology. There has been a tremendous progress in tailoring enzymes by protein engineering and design to develop "new and improved biocatalysts" (Bornscheuer et al., 2012; Schmid et al., 2001; Woodley, 2008). Various reports on improvement of biocatalysts include evolution of cellulases for applications in paper and textile industry (Datta et al., 2010; Kuhad et al., 2011; Lehmann et al., 2012), proteases for applications in detergent and laundry industry (Gupta et al., 2002; Martinez et al., 2013; Maurer, 2004), cytochrome P450s for applications in production of fine chemicals and pharmaceutical intermediates (Dennig et al., 2012; Urlacher and Girhard, 2012).

1.2 Cytochrome P450 Monooxygenases

Cytochrome P450 monooxygenase (CYP450s or CPYs) is one of the largest superfamily of heme-thiolate proteins. Present nearly in all genomes of the biological kingdom (De Montellano, 2005; Denisov et al., 2005), they catalyze oxidation of diverse range of endogenous and exogenous compounds in human, bacteria and plants (Bernhardt, 2006).

The terminology P450 describes the spectral properties of this enzyme class. Cytochrome P450s are named after their unusual spectral properties displaying absorption maxima at 450 nm of the reduced carbon monoxide bound complex (Omura and Sato, 1964). This spectral feature of P450s is induced by a cysteine thiolate group forming the fifth ligand of the heme iron and classifies cytochrome P450 enzymes as heme-thiolate proteins (Ichikawa and Yamano, 1967; Murakami and Mason, 1967). Cytochrome P450s usually act as terminal monooxygenases. The classification of P450s in the enzyme group is as shown in Figure 1.1.
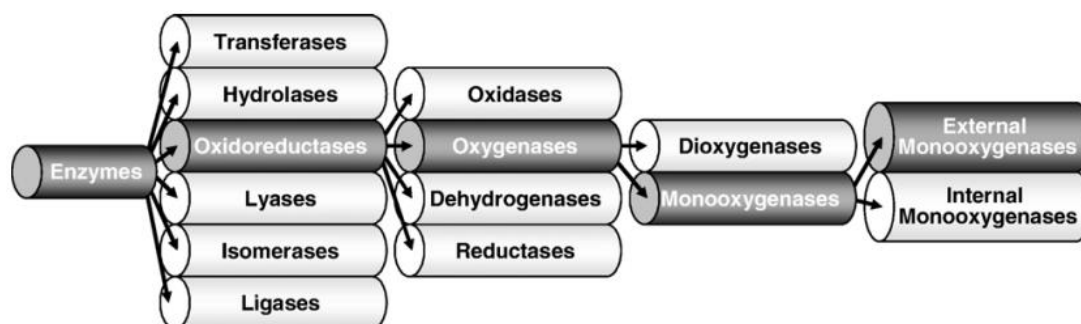
Figure 1.1. Assignment of Cytochrome P450 to enzyme group. P450 are highlighted with dark grey colored subgroups. Figure taken from (Hannemann et al., 2007).

Cytochrome P450 primarily catalyzes oxidation reaction by insertion of one atom of molecular oxygen into an organic substrate. Oxidation reactions are manifested as hydroxylation, oxygenation, epoxidation, dealkylation and are usually highly regioselective and stereoselective (Jung et al., 2011) (Figure 1.2). Selective C-H oxyfunctionalization at unactivated carbons, one of the most challenging reactions in synthetic chemistry is performed by P450s especially in water, at room temperature and under atmospheric pressure (Cryle et al., 2003; Whitehouse et al., 2012).
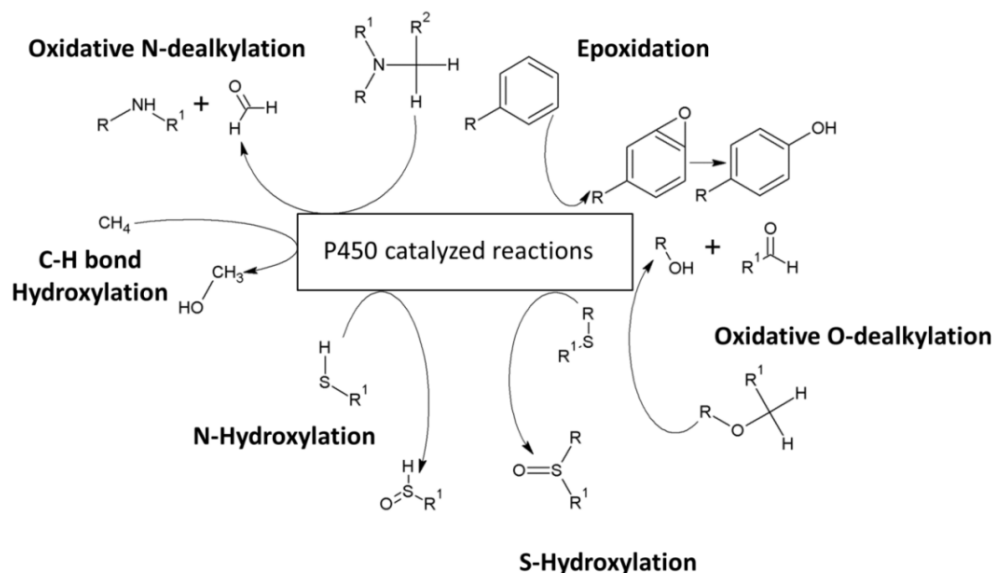
Figure 1.2. Reactions catalyzed by Cytochrome P450 monooxygenases (Cryle et al., 2003; Whitehouse et al., 2012; Fasan 2013).

1.2.1 Catalytic Mechanism of P450s

P450s catalyze the incorporation of one atom of molecular oxygen into various substrates. For this reaction P450s obtain their electrons from NADPH or NADP and shuttle it to the catalytic heme center via one or more electron transfer proteins (Jung et al., 2011; Urlacher and Girhard, 2012).

$$R\text{-}H + NAD(P)H + H^+ \xrightarrow[O_2]{P450 \text{ monooxygenase}} ROH + NAD(P)^+ + H_2O$$
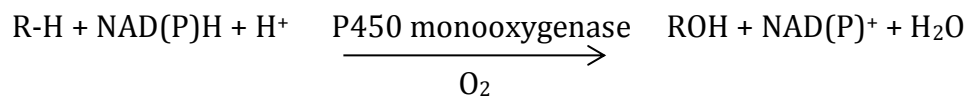
Scheme 1. A typical reaction catalyzed by P450 monooxygenases during hydroxylation of substrates. The reaction is characterized by insertion of one oxygen atom from molecular oxygen to the substrate (R-H) to form the product (R-OH) with parallel reduction of the other oxygen atom to water.

The catalytic cycle of P450s has been a topic of extensive research since 1950s in an attempt to decipher the mechanism of oxygen activation and its subsequent transfer to non-activated carbon atoms by P450 monooxygenases. The general scheme of reaction mechanism considered universal for all P450s is shown in Figure 1.3.

The substrate binds to the ferric form of protein in close proximity to the heme (Figure 1.3, step I). This results in a dramatic increase in redox potential of the heme iron, enabling it to accept an electron from the electron transfer protein. Transfer of electron is coordinated by electron transport proteins. The main function of these electron transfer proteins is uncoupling two electrons from the reducing equivalents (NADH or NADPH) and funneling them singly to the catalytic heme domain. The transfer of this electron results in the formation of ferrous state of the enzyme (Figure 1.3, step II). This reduced form of the enzyme can now bind to molecular oxygen forming ferrous dioxy complex (Figure 1.3, step III) which in turn is further reduced to ferric peroxy anion (Figure 1.3, step IV). Protonation of the ferric peroxy anion produces a highly unstable hydroperoxo-iron species $(Fe(III)-OOH)$ (Figure 1.3, step V). Hydroperoxo-iron species upon further protonation forms iron-oxo $Fe(V)=O$ complex (compound I) with simultaneous release of a water molecule (Figure 1.3, step V). Interaction of the substrate with compound I through radical rebound mechanism leads to the formation of a hydroxylated product (Figure 1.3, step VI) (De Montellano, 2005; Denisov et al., 2005; Makris et al., 2005; Meunier et al., 2004).
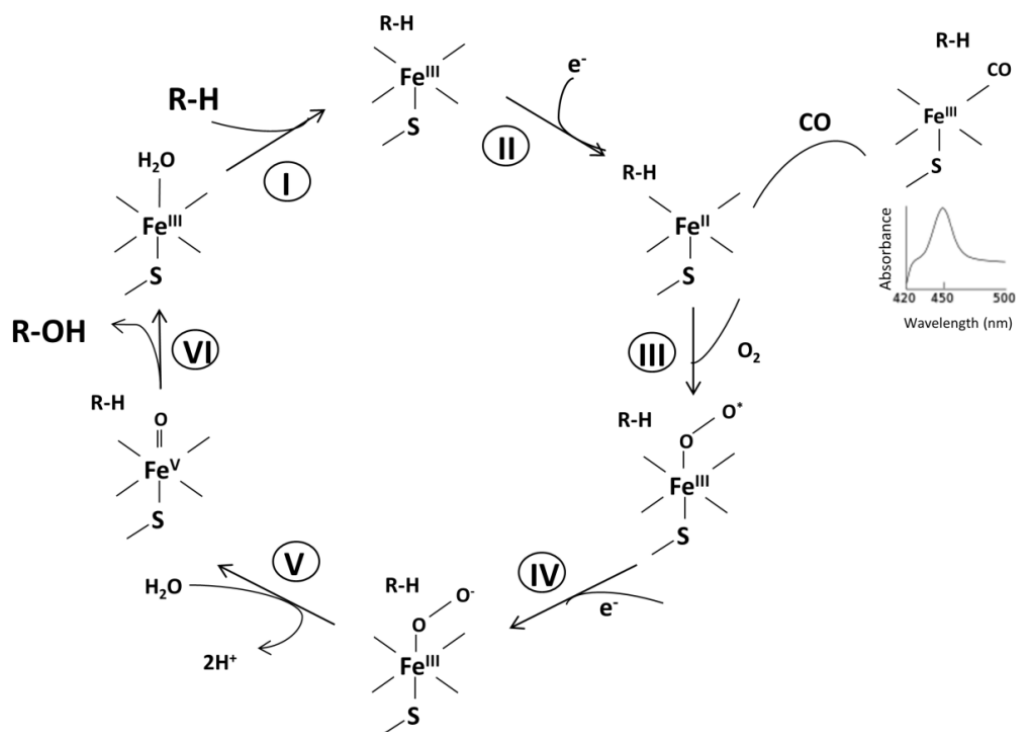
Figure 1.3. Catalytic cycle of P450 monooxygenases. Figure adapted from Cryle et al., 2003.

1.2.2 Classification of Electron Transport Systems for Cytochrome P450s

Cytochrome P450s belong to external monooxygenases (Figure 1.1), implicating the need for an external electron donor necessary for oxygen activation and subsequent substrate hydroxylation. Genomic, biochemical and biophysical characterization of cytochrome P450 monooxygenases revealed a great diversity in the molecular organization of the P450s and their redox partners (McLean et al., 2005; McLean et al., 2007; Urlacher and Girhard, 2012). Initially cytochrome P450s were grouped in two main classes. First, the adrenal mitochondrial P450 systems in which electrons were shuttled from NADPH via adrenodoxin and adrenodoxin reductase. Second, the liver microsomal P450s in which electrons from NADPH are transferred via a FAD-FMN containing reductase to the P450 monooxygenase. The first bacterial cytochrome P450 system discovered was a camphor monooxygenase and it was organized analogously to the mitochondrial one. Electrons were transferred via an iron –sulfur cluster [2Fe-2S] to P450cam. Later, CYP102 (P450 BM3) was discovered showing an unusual property, a fusion of the P450 and the reductase components into one polypeptide chain. Since then, several unusual electron transfer partner complexes for different P450s have been described (Figure 1.3, Table 1.1). Figure 1.4 and table 1.1 shows ten different classes of cytochrome P450s with their electron transfer chain components (Hannemann et al., 2007).
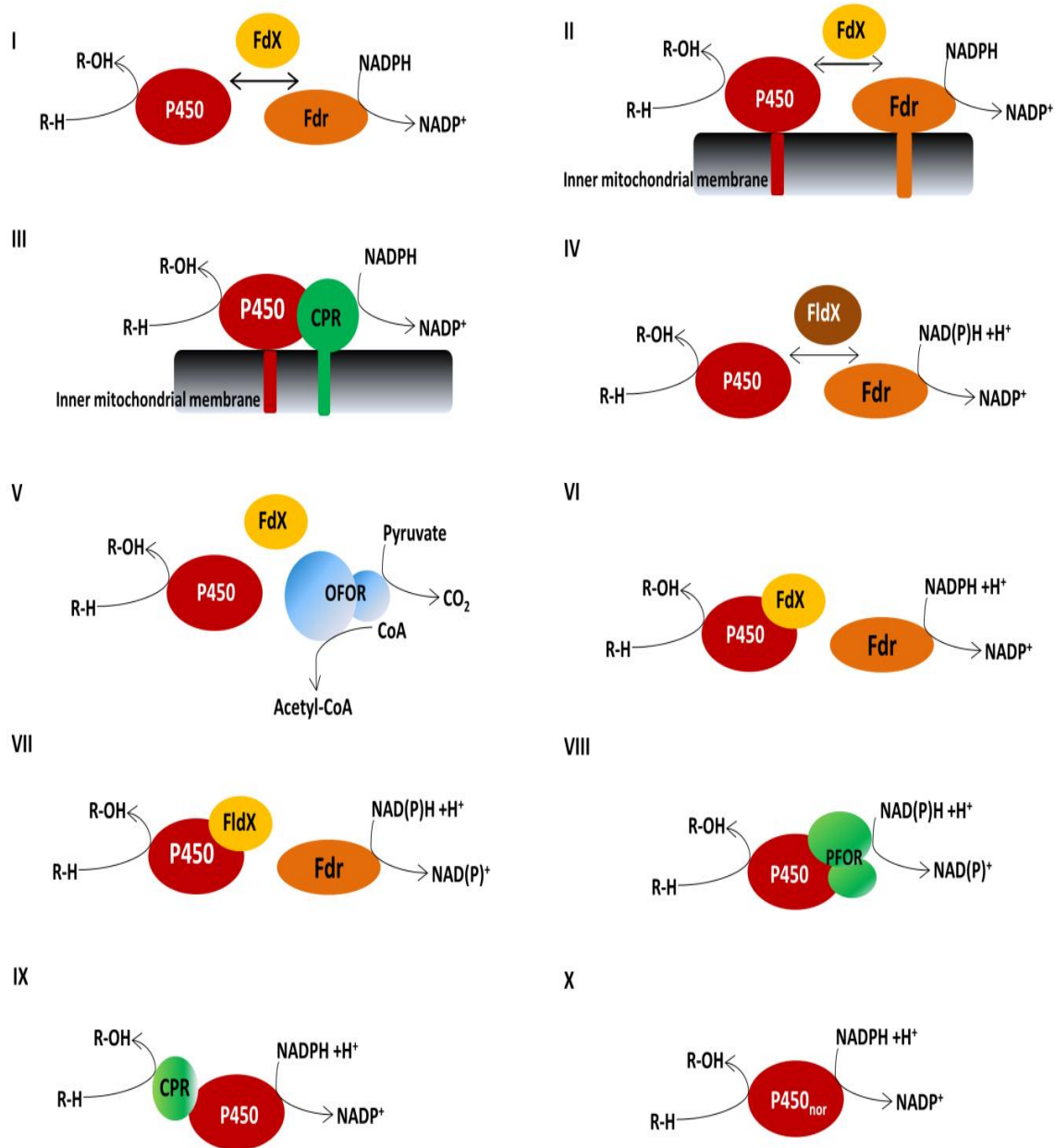
Figure 1.4. Organization of different cytochrome P450 systems. Adapted from (Hannemann et al., 2007).

(I) Class I, bacterial system; (II) Class II, mitochondrial system; (III) class II microsomal system; (IV) class III, bacterial system e.g. P450cin; (V) class IV, bacterial thermophilic system; (VI) class V, bacterial [Fdx]- [P450] fusion system; (VII) class VI, bacterial [Fldx]-[P450] fusion system; (VIII) class VII, bacterial [PFOR]- [P450] fusion system; (IX) class VIII, bacterial [CPR]- [P450] fusion system; (X) class IX, soluble eukaryotic P450nor.

Table 1.1. Classes of P450 systems classified depending on the topology of the protein components involved in the electron transfer to the P450 enzyme. Table has been taken from Hanemann et al., 2006.

Class/Source	Electron Transport Chain	Location/Remarks
Class I		Cytosolic, soluble, P450cam from
Bacterial	NAD(P)H → [Fdr] → [FdX] → [P450]	*Pseudomonas putida*
Mitochondrial	NADPH → [Fdr] → [FdX] → P450	
Class II		Cytosolic, soluble; *Streptomyces*
Bacterial	NADH → [CPR] → [P450]	*carbophilus*
Microsomal A	NADPH → [CPR] → [P450]	Membrane anchored, ER
Microsomal B	NADPH → [CPR] → [cytb5] → [P450]	Membrane anchored, ER
Microsomal C	NADH → [cytb5Red] → [cytb5] → [P450]	Membrane anchored, ER
Class III	NAD(P)H → [Fdr] → [FldX] → [P450]	Cytosolic, soluble, P450cin from
Bacterial		*Citrobacter braakii*
Class IV	Pyruvate → [OFOR] → [FdX] → [P450]	Cytosolic, soluble, *Sulfolobus*
Bacterial		*tokadaii*
Class V	NADH → [Fdr] → [FdX-P450]	Cytosolic, soluble, *Methylococcus*
Bacterial		*capsulatus*
Class VI	NAD(P)H → [FdR] → [Fldx-P450]	Cytosolic, soluble, *Rhodococcus*
Bacterial		*rhodochrous* strain 11Y
Class VII	NADH → [PFOR-P450]	Cytosolic, soluble, *Rhodococcus sp*
Bacterial		strain,
Class VIII	NADPH → [CPR-P450]	Cytosolic, soluble, *Bacillus*
Bacterial, Fungal		*megaterium*, *Fusarium oxysporum*
Class IX	[NADPH] → [P450]	Cytosolic, soluble, *Fusarium*
Only NADH		*oxysporum*
dependent, Fungal		
Class X	P450	Membrane bound, ER
Independent in		
plants /mammals		

1.2.3 Artificial Cytochrome P450 Fusion Proteins

Various classes of cytochrome P450s classified on the basis of proteins involved in electron transfer to heme center are described above (Table 1.1 and Figure 1.3). Most prokaryotic P450s are three component systems and eukaryotic P450s are membrane bound requiring another membrane associated redox protein for their function. A variety of polycistronic expression systems have been described for construction of catalytically active multicomponent systems systems (Kim and de Montellano, 2009; van Beilen et al., 2006).

However the application of multicomponent P450s as oxidation catalyst still remains challenging. After the discovery of natural fusion of heme and reductase domains (P450 BM 3), efforts were made to generate artificial fusion proteins or search for new self-sufficient cytochrome P450s. The first self-sufficient bacterial P450 system was constructed by fusing P450cam and its natural redox partners, putidaredoxin reductase (PdR) and putidaredoxin (Pdx); P450cam-Pdx-PdR catalyzed the conversion of camphor to 5-*exo*-hydroxy-camphor (Sibbesen et al., 1996). However, the activity of the fusion protein reached only 30% as compared to the activity obtained after using an equimolar mixture of its individual components. A recent study reports the construction of fusion protein fusion for P450cam by employing enzymatic cross linking of lysine tagged PdX with a peptide fragment between PdX and P450cam containing a reactive glutamyl residue prone to attack by transglutaminase. This type of fusion showed a 10-fold higher catalytic activity compared to the P450cam-Pdx-PdR fusion described by Sibbesen et al (Hirakawa et al., 2007). In another report, Hirakawa *et al.,* used PCNA (proliferating cell nuclear antigen), a trimeric DNA binding protein complex, to make a heterotrimeric P450. The three components of P450cam monooxygenase system were fused to PCNA1, PCNA2, and PCNA3 to generate PCNA1-PdR, PCNA2-Pdx, and PCNA3-P450cam, respectively. Although the specific activity of the fusion protein was decreased, the designed complex exhibited a 50-fold higher NADPH and oxygen consumption rates compared to the reconstituted system (Hirakawa and Nagamune, 2010).

Recently isolated native fusion of heme domain and its redox partner is P450RhF isolated from *Rhodococcus* sp. NCIMPB 9784. Sabbadin *et al.* described a general LICRED method for rapid generation of libraries of various heme domains fused to the reductase domain of P450RhF (RhFRED) as a general redox partner (Sabbadin et al., 2010). The LICRED method was demonstrated on heme domains from P450cam and P450 XpIA, and the resulting enzymes were shown to be active on their original substrates, d-camphor and hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX) (Sabbadin et al., 2010).

One important parameter that governs the activity of fusion proteins is the orientation of the fused protein domains and the length of the linker separating these domains. In case of P450 BM 3, naturally fused cytochrome P450, the C terminal of heme domain is fused to the N terminal of the reductase domain via short stretch of amino acid residues. The role of the

linker region is not clear however, it is commonly postulated that the linker region acts as a flexible hinge allowing recognition and interaction between the two domains (Hannemann et al., 2007; Sevrioukova et al., 1999). Furthermore, Munro *et al*. reported that the linker region might also be responsible for stabilization of the active dimeric form of P450 BM3. Previous reports demonstrate that the length of the linker is decisive for activity of fusion proteins rather than the amino acid composition of the linker. In two individual studies Govindraj and Poulos reported that changing Arg-Lys-Lys stretch in the middle of the linker sequence in P450 BM3 to Ala-Ala-Ala does not alter the functional properties of either of the two domains. However, deleting three or six amino acids leads to drastic effect of the hydroxylation activity of enzyme. The three amino acid deletion mutant exhibited 10% of the hydroxylation activity while the six amino acid deletion mutant showed nearly undetectable levels of fatty acid hydroxylase activity (Govindaraj and Poulos, 1995, 1996).

1.2.4 Applications of P450s

P450s have received a great attention from industry with an aim to exploit their reaction chemistry for production of fine chemicals. P450s also have potential application in bioremediation.

A well-established commercial application of P450s in 11β-hydroxylation is of Reichstein S to hydrocortisone with *Curvularis sp.* at a scale of ~ 100 tons/year: a process which allows regio and stereoselective hydroxylation by incorporation of oxygen in Reichstein S (www.schering.de; (Petzoldt K, 1982). Furthermore, it has also been used for the conversion of progesterone to cortisone by *Rhizopus sp*. (www.pfizer.com); (Peterson 1952; Hogg 1992).

Development of new and cheap resources for large scale usage of P450s will further aid in use of P450s for production of pharmaceuticals and fine chemicals. *Streptomyces* sp. strain Y-110 has been reported to produce Pravastatin, a cholesterol lowering drug (1 g/l) (Lin et al., 2011). Another example is *Saccharopolyspora erythraea* which catalyzes a hydroxylation step in the production of macrolide antibiotic erythromycin. The production of anticancer drug perrillyl alcohol from limonene has been reported using cytochrome P450 alkane hydroxylase isolated from *Mycobacterium* sp.

In addition to this, P450s have also been reported to be used for bioremediation purposes. Numerous reports discuss the biodegradation of environmentally harmful pesticides; insecticides are other agrochemicals (Guengerich, 1995; Kellner et al., 1997). Several P450 BM 3 mutants constructed using site directed mutagenesis are reported for degradation of polycyclic aromatic hydrocarbons such as phenanthrene, fluoranthrene and pyrene (Carmichael and Wong, 2001).

1.3 Cytochrome P450cin

Cytochrome P450cin (CYP176A1) is a bacterial monoxygenase isolated from *Citrobacter braakii* which can grow on 1,8-cineole as its sole source of carbon (Hawkes et al., 2002). P450cin catalyzes the stereoselective hydroxylation of 1,8-cineole to (2S)-2β-hydroxy 1,8-cineole or (S)-6β-hydroxycineole (Scheme 2) (Hawkes et al., 2002; Meharenna et al., 2004).

1,8-cineole   2β-hydroxy 1,8-cineole
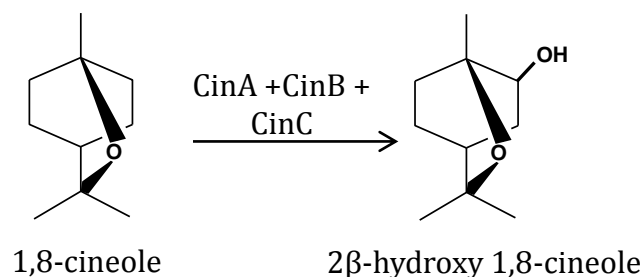
Scheme 2. Hydroxylation of 1,8-cineole to 2β-hydroxy 1,8-cineole with multicomponent Cytochrome P450cin.

1,8-cineole, a cyclic monoterpene ether, also known as eucalyptol, is the major component of the essential oils derived from *Eucalyptus polybractea.* It imparts the pleasant and unique odor associated with the eucalyptus derived essential oils (Miyazawa et al., 2001). In industry, it is widely used in cosmetic and pharmaceutical preparations for external applications in nasal sprays, disinfectants, as analgesic and food flavoring (Rodríguez et al., 2006). The hydroxylated products of 1,8-cineole are chiral intermediates for organic chemistry applications. In addition, 2-endo-hydroxy-1,8-cineole and its esters have antimicrobial and bactericidal activities and are also used as valuable aroma components (Miyazawa and Hashimoto, 2002; Miyazawa et al., 2001).

The three component electron transfer system of P450 consists of a heme-containing active center (P450cin/CinA), a flavin adenine dinucleotide (FAD)-containing cindoxin reductase (CinB), and a flavin mononucleotide (FMN)-containing cindoxin (CinC), encoded by the *cinA, cinB and cinC* genes respectively. The CIN operon consists of four open reading frames and are shown in figure 1.5 (Hawkes et al., 2002).
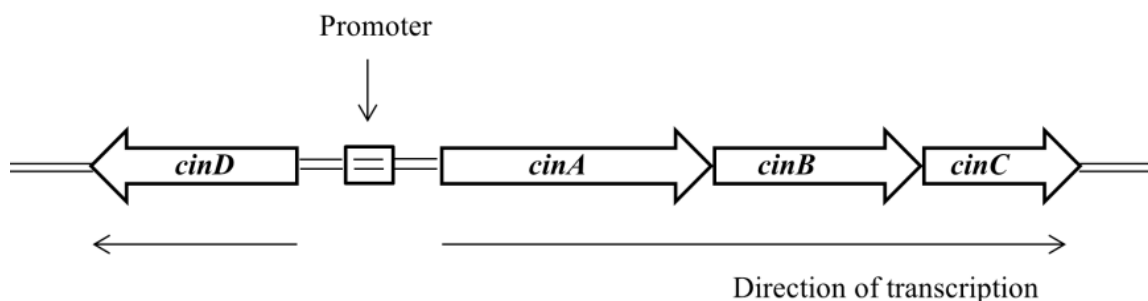
Figure 1.5. Map of the CIN operon that contains the four open reading frames, *cinA, cinB, cinC* and *cinD.*

The first ORF (*cinA*) encodes for a 405 amino acid protein with a molecular weight of 45 kDa. Homology modelling indicated that it belonged to P450 family being similar to a number of bacterial P450s. The second ORF (*cinB)* encodes a 49 kDa, 452 amino acid protein which is a NADPH dependent ferrodoxin reductase on the basis of sequence homology. *cinC* encodes a protein (16 kDa, 155 amino acids) that was most similar to the FMN containing domain of cytochrome P450 reductase based on homology modelling. A fourth ORF (*cinD)* encodes the protein (1R) - 6β-hydroxycineole dehydrogenase which displays a significant homology to a number of short chain alcohol dehydrogenases. The *cinD* gene is divergently transcribed with respect to *cinA, cinB and cinC* (Figure 1.5) (Hawkes et al., 2002; Slessor et al., 2010). *cinD* exhibits (1R)-6β-hydroxycineole dehydrogenase activity, the second step in the degradation of 1,8-cineole. (1R)-6β-hydroxycineole dehydrogenase displays a strict requirement for NADH, with no reaction observed in the presence of NADPH. The enzyme also catalyzes the reverse reaction, reducing (1R)-6-ketocineole to (1R)-6β-hydroxycineole (Figure 1.6) (Slessor et al., 2010). P450cin shares many properties with P450cam (27 % identity and 46 % similarity) and other bacterial P450s however has two unique features. First, the highly conserved,

catalytically important threonine residue in the distal I-helix in all P450s involved in dioxygen activation is replaced by the Asn242 in P450cin. Site directed mutagenesis and crystallographic studies have suggested that Asn242 does not appear to be critical for O_2 activation but instead is essential in regio-selective hydroxylation of the substrate. The second difference is unlike P450cam, the active site is not surrounded by non-polar residues but instead an intricate lattice of water bridged hydrogen bonds. This water network involves residues Tyr81, Asn242 and Asp241.

1,8-cineole   (1R)-6β-hydroxycineole   (1R)-6-ketocineole
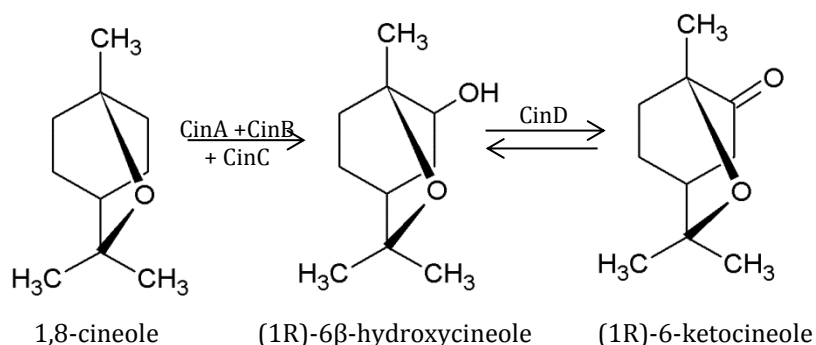
Figure 1.6. Proposed pathway for 1,8-cineole degradation. The first step is the hydroxylation of 1,8-cineole to (1R)-6β-hydroxycineole, which is further oxidized to (1R)-6-ketocineole by *cinD*.

Another unique feature of cytochrome P450cin electron transfer system is the presence of multicomponent reductase system comprising of FAD (cindoxin reductase, CinB) and FMN (cindoxin, CinC) containing electron transfer proteins (Slessor et al., 2010). CinB has strong similarity with the FAD containing ferredoxin reductase found in class I adrenal mitochondrial system whereas CinC has a strong similarity with class II reductase.

1.4 Electron Transfer in Redox Proteins

Redox (reduction-oxidation) reactions include all chemical reactions in which atom(s) have their oxidation state changed and are characterized by transfer of electrons between different species. The transfer of electrons between and within proteins is an essential feature for many physiological processes including biological energy transfer, metabolism

and enzymatic catalysis (Cordes and Giese, 2009; Gray and Winkler, 2009, 2010). The electron transfer process in redox protein is complex, versatile and tightly regulated in order to minimize energy loses and production of toxic metabolites like hydrogen peroxide and superoxide anions (Saab-Rincón and Valderrama, 2009; Wong and Schwaneberg, 2003).

Specific and efficient electron transfer in redox proteins depends on the interaction of two factors (i) the amino acid residues of the redox enzyme and (ii) the distribution of charged and uncharged amino acid residues. The primary sequence of amino acids determines the conformation for the co-factor binding pocket which in turn might determine the interaction of co-factor and amino acid residues necessary for electron transfer. Secondly, the native fold pattern of a protein displays a unique distribution of charged and uncharged amino acid side chains that determine the electrostatic potential surrounding the molecule (Saab-Rincón and Valderrama, 2009).

Redox enzymes generally contain metal ions like iron, copper, manganese and are arranged in order to assists electron flow and charge separation. The redox centers are placed in order of their redox potential, so that the direction of electron flow is established thermodynamically (Bizzarri and Cannistraro, 2005). The process of electron transfer in redox proteins can be well explained by Marcus theory. In Marcus theory, three factors govern the electron transfer within proteins: reorganization energies (quantitatively reflecting the structural rigidity in the oxidized and reduced form); potential differences and orientations of involved redox active sites; and the distance between the redox active site and intervening media (Marcus and Sutin, 1985).

Direct electrical communication between redox active center and electron donors is usually prohibited. Redox enzymes and electrode can be regarded as an electron donor-acceptor pair. Electron transfer theory defines the rate of electron transfer between a donor acceptor pair by equation 1,

$$k_{et} \propto e^{-\beta(d-d0)]}e^{\frac{-(\Delta G^0 + \lambda)^2}{4RT\lambda}} \qquad \text{Equation 1}$$

Where ΔG^o is the change in free energy associated with electron transfer and the generation of the redox products; λ is the reorganization energy; d^0 is the van-der-Waals distance; d the distance separating the donor and acceptor; and β is the electronic coupling co-efficient.

The mechanism of electron transfer in cytochrome P450s usually involves protein-protein interaction where electrons are transferred from redox proteins to the P450s in a complex electron transfer chain (Saab-Rincón and Valderrama, 2009). The redox centers are often buried in the protein core, with the protein serving as an insulator and a tight regulation over the electron transfer mechanism to avoid wastage of energy resources (Wong and Schwaneberg, 2003). One way of improving the activity of redox proteins is by protein engineering techniques particularly using rational design. The strategies to engineer a redox protein include trimming the protein, protein surface modification, active site mutation, domain shuffling or using alternative cofactor systems to substitute or replace NADPH in biocatalytic conversions.

1.5 Alternative Co-factor Systems for Cytochrome P450s

The use of isolated cytochrome P450 monooxygenase in catalysis is restricted due to the continuous requirement of NADPH in equimolar concentration during technical applications. NADPH is extremely expensive to use in cell free reaction systems, decomposes over time and is difficult to recover once oxidized (Udit and Gray, 2005). Therefore several approaches to substitute or recycle NADPH during biocatalytic conversions have been explored. These include co-factor regeneration using enzymatic or electrochemical approaches or co-factor substitution systems like mediated electron transfer, light induced systems, peroxide shunt pathway, electrochemically driven reactions. A useful co-factor regeneration/substitution method should fulfil several requirements. First, the enzymes, reagents, electrodes, immobilization matrices, should be readily available, inexpensive, easy to manipulate and stable under operational conditions. Second reagents used or byproducts of regeneration/substitution step should not negatively affect product stability and isolation (Holtmann and Schrader, 2007).

1.5.1 Co-factor Regeneration

Co-factor regeneration is one of the most commonly used strategy for whole cell biocatalysis using cytochrome P450s. Enzymatic co-factor regeneration is most commonly reported using enzymes formate dehydrogenase (FDH), glucose dehydrogenase (GDH),

glucose-6-phosphate dehydrogenase (G6PDH), and alcohol dehydrogenase (ADH) (Kubo et al., 2006; Maurer et al., 2005; Maurer et al., 2003; Michizoe et al., 2005; Müller et al., 2013; Wandrey, 2004). But, enzymatic co-factor regeneration approaches generally suffer from low enzyme stabilities and complexity of product isolations (Wichmann and Vasic-Racki, 2005). The most common system is using a dehydrogenase along with a substrate which can be oxidized with simultaneous regeneration of NAD(P)H from NADP$^+$ (for instance oxidation of formic acid to CO_2 using formate dehydrogenase)(Tishkov et al., 1999). Limitation of this system is the presence of two enzymes in the reaction system and most dehydrogenases use NAD$^+$ contrary to NAD(P)H like most P450s. To tackle this, P450 monooxygenases have been engineered to use NADH as cofactor instead of NADPH (Müller et al., 2013). NADH has the advantage to be less expensive and more stable than NADPH. Nevertheless, *in vivo* recycling systems suffer from drawbacks such as limited substrate uptake, toxicity of the substrate or product, product degradation and difficult downstream processing (Hollmann et al., 2006) as a result various co-factor substitutions systems have been reported as described below.

1.5.2 Co-factor Substitution

Strategies to substitute NAD(P)H include use of light induced systems (Shumyantseva et al., 2002), use of peroxide shunt pathway (Cirino and Arnold, 2003; Joo et al., 1999b), using mediators (Nazor et al., 2008; Nazor and Schwaneberg, 2006; Prasad et al., 2005; Schwaneberg et al., 2000; Shumyantseva et al., 2000).

In case of light induced systems, the catalytic enzyme complex is directly reduced by light. Shumyantseva et al., reported the covalent attachment of riboflavin to the mammalian cytochrome CYP2B4, yielding an artificial flavocytochrome for photo-induced intermolecular electron transfer between the isoalloxazine ring cycle of flavins and the ferric heme group of cytochrome CYP2B4 (Shumyantseva et al., 2002). Electrochemical regeneration of NADPH was described by Reipa et al., through its natural redox partner (putidaredoxin) using antimony doped tin oxide working electrode. The oxygen required for thr reaction was produced at a Pt counter electrode by water electrolysis. A continuous catalytic cycle continued for more than 5 h with approximately 2,600 enzyme turnovers. The maximum product formation rate was 36 nmol of 5-*exo* hydroxycamphor/nmol of

CYP101 per min (Reipa et al., 1997). The use of peroxide shunt pathway for P450 catalyzed reactions is an interesting approach, however, the enzyme rapidly inactivates at peroxide concentration required for catalysis (Cirino and Arnold, 2003; Joo et al., 1999b). Alternative approaches include the direct electron transfer (Nazor and Schwaneberg, 2006; Schwaneberg et al., 2000) and mediated electron transfer (Nazor et al., 2008) as described below.

1.5.2.1 Direct Electron Transfer Systems

Direct electron transfer involves the direct electron transfer to the active site of the enzyme from an electron donor. Direct electron transfer was reported for P450 BM 3 in which electrons were directly transferred from a conducting polymer Baytron into the active site of P450 BM 3 (Nazor, 2007). However the electron transfer is not yet understood on a molecular level.

Direct electron transfer using electrochemical approach has been reported using a variety of metal and non-metal electrodes but often without any proof of electrochemically driven substrate conversion. One major disadvantage of using direct electrochemically mediated catalysis is the instability of enzymes upon interaction with the electrode surface (Holtmann and Schrader, 2007).

Protein engineering and protein surface modification have been reported to tackle this problem. Ferapontova and Gorton reported a six-histidine tag at the C terminus of horseradish peroxidase, which was adsorbed on a pre-oxidized gold electrode and generated more than 30-fold increase in electron transfer rate (Tosstorff et al., 2014). Another example was addition of a polylysine chain at the C terminus of glucose oxidase. The polylysine chain was added to anchor the electron transfer mediator ferrocenecarboxylic acid, for improving sensitivity and stability of glucose biosensors. The modified glucose oxidase showed comparable kinetic parameters to those of the wild-type however, the polylysine tagged enzyme retained 90% of its activity whereas the wild type retained only 20% (Ferapontova et al., 2002). Another strategy which was reported for improving the direct electron transfer include trimming or removing of amino acids residues around the active site which are not essential for maintaining protein structure

and function. This strategy was successfully employed for improving the electron transfer rates of microperoxidase and laccases (Petzoldt K, 1982; You and Arnold, 1996). Davidson reported the introduction of active site mutations in methylamine dehydrogenase which led to a fourfold improvement in the sensitivity of a histamine sensor owing to a lowered K_m value (Bao et al., 2002; Davidson, 2003).

An alternative to direct electron transfer approach is the mediated electron transfer systems. Mediated electron transfer systems make use of a mediator (a redox compound) which acts as a shuttle between the electron source and enzyme active site. It overcomes the limitations of direct electron transfer like dependency of the protein orientation on electrode surface or enzyme inactivation.

1.5.2.2 Mediated Electron Transfer Systems

Electroactive heme center is deeply buried inside the protein core and the transfer of electrons depends on the orientation of protein and the proximity of protein with the electron source. In such a case a mediator can be used to increase the electron transfer rates between the electron source and active site of the enzyme. In such a reaction set up, mediator must be able to be reduced, e.g. at an electrode at an appropriate potential, and the reduced form must be able to transfer the enzyme active site. Faulkner et al., first described Cobalt(III)sepulchrate as a mediator for P450 catalyzed hydroxylation reactions using platinum electrodes. They replaced NADPH for hydroxylation reaction using mediator Co(III)sepulchrate for the electro-catalytically driven ω-hydroxylation of lauric acid with P450 BM 3. However, the turnover rates were nearly 9 folds lower while using Co(III)sepulchrate as compared to those obtained for NAD(P)H (Faulkner et al., 1995). Cobalt sepulchrate has found a broad application as mediator in P450 catalyzed reactions. It acts as an one electron shuttle between the electron source and enzyme, where the cobalt atom switches between Co(II) and Co(III) oxidation states (Nazor and Schwaneberg, 2006). One promising approach for expensive cofactor regeneration or substitution is the use of electrochemical methods. Application of electrochemical methods with P450s has also attained a growing attention due to the possibility of developing applications such as biosensors and bioreactors (Holtmann and Schrader, 2007; Shumyantseva et al., 2005).

Schwaneberg et al., reported the use of Cobalt(III)sepulchrate as mediator for cytochrome P450 BM 3 in solution in combination with zinc dust as a cheap source of electrons. Using the Zn/Co(III)sepulchrate system a turnover rates of 125 nmol substrate consumed nmol^{-1} P450 was obtained which was considerably lower as compared to turnover rates of 574 nmol substrate consumed nmol^{-1} P450 using NADPH as a reduction equivalent for a mutant of P450 BM3 (Schwaneberg et al., 2000).

Although, the direct and mediated electron transfer systems provide a promising approach for substitution of NAD(P)H in cell free catalysis, they suffer from various drawbacks like enzyme inactivation, low enzyme stability, low turnover numbers and low turnover frequencies. These drawbacks can be tackled by developing biocatalysts using protein engineering techniques to make up for these inadequate properties to obtain a variant for electrobiocatalytic applications.

1.6 Protein Engineering

The application of enzymes is constantly increasing for synthesis of pharmaceutical intermediates and fine chemicals. The development of genetic engineering methods enabled an expansion on the use of enzymes in industry using protein engineering tools. Naturally occurring enzymes (wild type) are optimized for physiological conditions; however they are not suitable for industrial applications (Arnold, 2001). Limitations include low stability (towards temperature, pH and oxidative agents), narrow operating conditions (pH, temperature), poor activity in non-aqueous solvents, and low catalytic activity for non-natural substrates. The aim of protein engineering approach is to tailor and redesign the biocatalysts to make them suitable for industrial applications (Farinas et al., 2001).

It also generates valuable information to understand structure- function relationship of an enzyme class (Wong et al., 2007). It basically mimics the process of Darwinian evolution in a test tube. It however differs from the Darwinian evolution, firstly, it does not create new enzymes but improves the properties of existing one (microevolution) and secondly, the time scale is reduced from millions of years (natural evolution) to a few months or years (protein engineering) (Stemmer and Holland, 2003). Besides improving a native property of

an enzyme, protein engineering can also be used to impart and improve non-native properties in an enzyme. The most widely used approaches for protein engineering are rational design and directed evolution (Arnold, 2001; Wong et al., 2007). A comparison of different enzyme engineering approaches is given in Table 1.2.

Table 1.2. Comparison of different enzyme engineering approaches (Chica et al., 2005).

	Rational Design	Random Mutagenesis	Semi-Rational design
High throughput screening	Not essential	Essential	Advantageous
Structure and/or functional information	Both essential	Neither essential	Either sufficient
Probability of obtaining synergistic mutants	Moderate	Low	High
Sequence space exploration	Low	Moderate	Computational: vast Experimental:moderate

1.6.1 Directed Evolution of Enzymes

Directed protein evolution is an approach which was introduced in 1967 (Yamane et al 2004). Starting a directed evolution experiment does not require any knowledge of the target protein. It can be applied to enzymes for which no crystal structure is available and/or key amino acid positions are unknown. Major steps involved in a directed evolution experiment are presented in Figure 1.6. It comprises of **I)** generation of diversity, **II)** screening of generated mutant library in a screening condition comparable to the final application, **III)** isolation of gene encoding for the improved protein variant. This process is repeated in an iterative cycle until a protein variant having desired property is obtained (Wong et al., 2007).
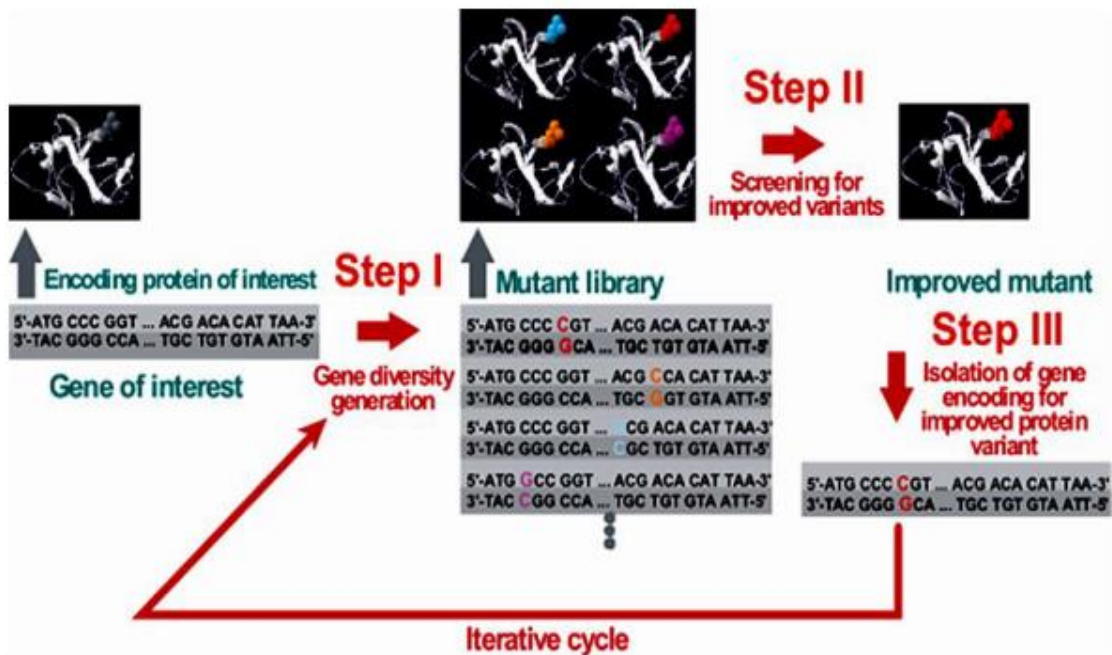
Figure 1.7. Scheme of a directed evolution experiment in an iterative cycle**. Step I**-generation of mutant library, **step II**- screening of the generated mutant library to obtain an improved variant, **step III**- isolation of gene encoding for the improved protein variant. Figure taken from (Güven et al., 2010).

The two main challenges in a directed evolution campaign are unbiased diversity generation and developing a screening/selection system that mimics the desired final working condition (Schmidt-Dannert and Arnold, 1999). A variety of chemical and molecular biology methods are available for generation of random mutagenesis libraries in the gene of interest. A perfect diversity generation method should fulfill seven criteria: I) unbiased mutational spectrum II) controllable mutation frequency, III) consecutive nucleotide substitution or codon-based substitution, IV) enable subset mutagenesis, V) independence of gene length, VI) technically simple and reproducible and VII) economical (Wong et al., 2007).

Methods to introduce random mutations include PCR based random mutagenesis methods (Cadwell and Joyce, 1994), using mutator strains, using chemical mutagens (Mohan and Banerjee, 2008), UV radiation (Cox and Parry, 1968) or recombination methods like DNA shuffling (Stemmer, 1994), Staggered Extension Process (StEP) (Zhao et al., 1998). The most commonly used are PCR based method under low fidelity conditions. PCR based

methods are based on the use of error prone polymerase or biased nucleotide or Mn^{+2} concentrations. PCR based methods for generation of random libraries are transition biased, which fail to generate consecutive mutations and as a result fail to cover the entire sequence space. A recently developed method which overcomes the limitations of PCR based methods is Sequence Saturation Mutagenesis (SeSaM) (Wong et al., 2004). SeSaM enables introduction of mutations by randomizing each amino acid position in the protein sequence. It also overcomes the limitation of polymerase bias (transitions over transversions) and can introduce subsequent mutations. Thus SeSaM generates a higher theoretical diversity that potentially increases the likelihood of finding improved variants. The method used for generation of a random library depends on the property of the enzyme which is to be improved and also the screening capabilities. For instance, engineering an enzyme for thermostability requires high mutational load and substitution with chemically different amino acids making SeSaM the method of choice for diversity generation. While, engineering an enzyme for enantioselectivity often requires substitution with similar kind of amino acids making epPCR, the method of choice (Arnold, 2001; Reetz and Jaeger, 2000). However, only the generation of a high quality mutagenic library does not suffice the need of a directed evolution experiment but a reliable and high throughput screening system is necessary to screen the generated mutant library in order to obtain improved variant(s). The screening system should as close as possible be able to replicate the conditions for final application (Aharoni et al., 2005). The most common screening methods are based on solid phase screening assays and microtiter plate assays. Solid phase screening can be performed directly on an agar plate, on a filter paper or on a membrane. These types of screening systems have a high throughput (10^4 to 10^6) and are based on visual signals like color development or formation of halos. However, they usually have low accuracy and are generally used as pre-screens for differentiation between active and inactive clones (Lin and Cornish, 2002)

Screening systems centered on microtiter plates (96-, 384- or 1536 well) are usually based on colorimetric or fluorometric reactions and are currently the most widely used screening technique with a throughput of 10^3 to 10^5. They can be performed using cell lysates or supernatants and enzyme activity is measured using a spectrophotometer or a fluorometer (Aharoni et al., 2005). The introduction of flow cytometry based screening systems in

directed evolution enable the screening of mutant libraries with more than 10^7 clones per hour (Farinas, 2006). This system is based on detection of fluorescent product generated by enzymatic reaction in which the enzyme is linked to the cell surface or entrapped in a double emulsion (Miller et al., 2006). The use of FACS (fluorescence activated cell sorter) and microfluidic devices greatly accelerates the screening system, however currently it offers only the qualitative differentiation of positive and negative variants. In addition, dye diffusion and limited number of fluorescent based assays for enzymes which are necessary for flow cytometric screening are still the prevailing challenges. Table 1.3 summarizes the comparison of currently available screening formats for screening of mutant libraries in a directed evolution experiment. A typical random library usually consists of 10^6-10^8 clones required to cover a considerable portion of protein sequence space. Thus a high throughput, reliable screening system is required to screen such a large amount of clones in order to select improved variant(s).

Table 1.3. Comparison of currently used screening techniques in directed evolution. Table reproduced from (Leemhuis et al., 2009).

Strategy	Library Size	Advantge	Disadvantage
Selection	$\sim 10^9$	Yields desirable variants only	Only possible if property confers an advantage
Agar plate screening	$\sim 10^5$	Simple to operate	Limited dynamic range
Microtiter plate screening	$\sim 10^4$	Most analytical methods possible. Excellent dynamic range	Relative low screening capacity
Cell in droplet method	$\sim 10^9$	Large libraries	Fluorescence detection and DNA modifying enzymes
Cell as microreactor	$\sim 10^9$	Large libraries	Fluorescence detection
Cell surface display	$\sim 10^9$	Large libraries	Fluorescence detection
In vitro compartmentalization	$\sim 10^9$	No cloning steps. Large libraries	Fluorescence detection

1.6.2 Rational Evolution of Enzymes

Rational evolution of enzymes target one or more preselected amino acid residue(s) in the target gene. The selection of the amino acids residue(s) is based on previous structural and

functional knowledge of the enzyme or crystal structure of the enzyme to identify promising residues (Chica et al., 2005). The selected amino acid is exchanged to a specific amino acid e.g. by site directed mutagenesis (SDM) or saturated using site saturation mutagenesis. Prediction of the 'best' amino acid at a particular position is difficult and often a site saturation mutagenesis method is applied. SSM allows the identification of the best amino acid on a specific position by applying a NNN (64 codons) or NNK degeneracy (32 codons), covering all 20 canonical amino acids (Reetz et al., 2008). Screening of focused library can be generally achieved using a medium throughput assay in a MTP. Rational design not only requires the knowledge of the crystal structure of enzyme but also of the property which should be improved. Many properties are well understood for instance thermostability of an enzyme can be improved by introduction of disulfide bridges or salt bridges which increase the rigidity of the protein scaffold (Korkegian et al., 2005). The experiments based on rational evolution are significantly less time consuming and accelerates the generation of new biocatalysts by generation of small and high quality mutant libraries (Reetz et al., 2008; Wong et al., 2007).

2. MATERIALS

Instruments, kits and devices were used according to supplier's recommendation.

2.1 Chemicals

Chemicals used were of analytical grade or higher quality and purchased from Sigma Aldrich Chemie (Taufkrichen, Germany), AppliChem (Darmstadt, Germany) or Carl Roth (Karlsruhe, Germany). Chemicals purchased from other suppliers or synthesized in-house are mentioned below:

Cobalt (III) sepulchrate	Synthesized as previously reported (Gahan, 1989)
Sodium hydrosulfite	ABCR GmbH, Karlsruhe, Germany
Yeast Extract	Merck KGaA, Darmstadt, Germany

2.2 Instruments and Devices

Äktaprime plus with UV detector	GE Healthcare, München, Germany
Autolab Potentiostat/Galvanostat	Metrohm, Riverview, USA.
Bench top shaker Certomat II	Sartorius, Göttingen, Germany
BioPhotometer Plus	Eppendorf, Hamburg, Germany
Centrifuge 5424, 5810R	Eppendorf, Hamburg, Germany
CO-lecture bottle station	Sigma Aldrich, Steinheim, Germany
Deep well plates (96 well-round bottom)	Brand GmbH, Wertheim, Germany
Deep well plates (96 well-V bottom)	Eppendorf, Hamburg, Germany
Evaporator IKA RV 10 basic	IKA-Werke, Staufen, Germany
GC-FID 2010 Plus	Shimadzu GmbH, Duisburg, Germany
Gel documentation U: Genius	Syngene, Cambridge, UK
Glass columns TAC15/125PE5- AB-2	YMC Europe GmbH, Dinslaken, Germany
Heat Gun	Steinel HL, Aachen, Germany
High Pressure Homogenizer	Avestin, Ottawa ON, Canada
Lyophilizer Alpha 1-2 LD plus	Christ, Osterode am Harz, Germany
Luggin capillary	Seele-Glas, Swisttal-Strassfeld, Germany

Polypropylene Microtiter plates (96 well)   Eppendorf, Hamburg, Germany

Microtiter plate shaker Microtron   INFORS, Heinsbach, Germany

Microtiter plate reader Tecan Sunrise   Tecan Group AG, Männendorf, Switzerland

Mini-Protean Tetra Cell system   Bio-Rad, München, Germany

Multipoint magnet stirrer Variomag 15   Thermo Fischer Scientific,Wilmington USA

NanoDrop Spectrophotometer 1000   Thermo Fischer Scientific,Wilmington USA

PCR thermocycler Mastercycler pro S   Eppendorf, Hamburg, Germany

Platinum   Allgemeine, Pforzheim, Germany

Pre-coated Silica gel 60 F254 plates   Merck KGaA, Darmstadt, Germany

Resin dispenser 48 position   Mettler- Toledo, Columbus OH, USA

Silver-Silver Chloride electrode   SensortechnikGmbH,Kaufbeuren, Germany

Sonicator Vibra-Cell VCX130   Sonics & Materials, Newton, USA

Spectrophotometer Varian Cary 50 UV   AgilentTechnologies, Darmstadt, Germany

Thermomixer Comfort   Eppendorf, Hamburg, Germany

Thin layer chromatography Unit*   CAMAG, Muttenz, Switzerland

Vortex Genie 2   Scientific Industries Inc., New York, USA

2.3 Cultivation and Expression Media

All media were sterilized by autoclaving (20 psi pressure, 121 °C, 20 min) and were stored at room temperature or frozen as 1 mL aliquots at -20 °C (SOC media) until further use.

Lysogenic Broth (LB medium) contained 10 g L^{-1} peptone, 5 g L^{-1} yeast extract, 10 g L^{-1} NaCl; pH 7.5 adjusted with 1 M HCl or NaOH. For preparation of agar plates the medium additionally contained 20 g L^{-1} agar-agar.

Terrific Broth (TB medium) contained 12 g L^{-1} peptone, 24 g L^{-1} yeast extract, 4 g L^{-1} glycerol (solution A; 800 mL), 2.31 g L^{-1} KH_2PO_4 and 12.54 g L^{-1} K_2HPO_4 (solution B; 200 mL). Solutions A and B were prepared separately and mixed together after autoclaving.

---

* Thin layer chromatography unit consists of Automatic TLC sampler 4 (ATS 4), Automatic developing chamber (ADC 2) and TLC scanner 3

Super Optimal Broth with catabolite repression (SOC medium) contained 20 g L^{-1} tryptone, 2 g L^{-1} NaCl, 5 g L^{-1} yeast extract and 2.5 g L^{-1} KCl. Sterile solutions of 10 mM $MgCl_2$, 10 mM $MgSO_4$ and 20 mM glucose were added after autoclaving.

2.4 Media Additives

All media additives were prepared as 1000X stock by dissolving individual components in MilliQ water and sterilized by filtration (0.2 µM *polyvinylidene difluoride* (PVDF) filter). Media additives were used in 1:1000 ratio in all media. Trace element solution was stored at 4 °C, magnesium chloride and sodium sulphate at room temperature until use. All other solutions were aliquoted (1 mL) and stored at -20°C until use.

Kanamycin sulphate solution contained kanamycin sulphate (100 mg mL^{-1}) dissolved in MilliQ water.

Trace element solution contained Calcium chloride dihydrate (0.45 mM), Zinc sulphate heptahydrate (0.65 mM), Manganese sulphate monohydrate (0.6 mM), Disodium EDTA (54 mM), Ferric chloride hexahydrate (62 mM), Copper sulphate pentahydrate (0.6 mM) and Cobalt chloride (0.75 mM) dissolved in MilliQ water.

Aminolevullinic acid solution (ALA) solution contained 5-Aminolevullinic acid (0.5 M) dissolved in MilliQ water.

Isopropyl-β-D-thiogalactopyranoside (IPTG) solution contained Isopropyl-β-D-thiogalactopyranoside (1 M) dissolved in MilliQ water.

Thiamine hydrochloride solution contained thiamine hydrochloride (0.5 M) dissolved in MilliQ water.

Sodium chloride solution contained sodium chloride (1 M) dissolved in MilliQ water.

Magnesium chloride solution contained Magnesium chloride (1 M) dissolved in MilliQ water.

2.5 Other Reagents, Solutions and Buffers

Iodine cleavage solution was prepared by mixing 5 µL of cleavage buffer (500 mM Tris, pH 9.0), 2 µL of iodine (100 mM in 99 % ethanol) in 2 µL MilliQ water (final volume of 10

μL). The iodine cleavage solution was prepared fresh and wrapped in aluminum foil and refrigerated (4 °C) until use.

Piranha solution contained conc. H_2SO_4: 30% H_2O_2 solution (3:1). The solution was prepared on ice under a fume hood cabinet by drop-wise addition of 30% H_2O_2 solution in conc. H_2SO_4 with gentle swirling.

Phosphate buffer was prepared by mixing equimolar (50 mM) solutions of dipotassium phosphate and monopotassium phosphate until the desired pH value (7.4 or 7.8) was achieved.

Tris/HCl buffer was used for purification of P450cin fusion protein. It contained 500 mM Tris/HCl, pH 7.8 (6.05 g L^{-1} Tris, pH adjusted by titration with 37% (v/v) HCl). Tris/HCl elution buffer contained additionally 1 M NaCl. The buffers were freshly prepared and filtered using a glass filter.

2.6 Strains, Vectors and Genes

The strains used in this work (Table 2.1) are *E. coli* DH5α for sub-cloning and *E. coli* BL21-Gold (DE3) lacI[Q1] for recombinant protein expression. The vectors used in the study are listed in Table 2.2.

Table 2.1. Genes employed in this work (see sequence in the appendix).

Gene	Origin	Reference
cinA	*Citrobacter braakii*	(Hawkes et al., 2002).
cinC	*Citrobacter braakii*	(Hawkes et al., 2002).
Fpr	*Escherichia coli*	(Jenkins and Waterman, 1994)

Table 2.2. Strains used for genetic manipulation and recombinant protein expression.

Strain	Genotype	Reference
E. coli DH5α	F- φ80lacZΔM15 Δ(lacZYA-argF)U169 recA1 endA1 hsdR17(rk-, mk+) phoA supE44 thi-1 gyrA96 relA1 λ-	Agilent Technologies
E. coli BL21-Gold (DE3)	F– ompT gal dcm lon hsdSB(rB- mB-) λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5])	Agilent Technologies
E. coli BL21-Gold (DE3) LacI^{Q1†}	Same as *E. coli* BL21Gold (DE3); additionally pGRG36-lacI^Q was used for site-specific integration of lacI into the chromosome	(Blanusa et al., 2010).

Table 2.3. Vectors used for genetic manipulation and recombinant protein expression.

Vector	Promoter	Antibiotic resistance/ Concentration	Replicon	Reference
pET28a(+)	T7	Kanamycin (Kan)/ 100 μg mL^{-1}	ColE1 (pBR32)	Novagen (Merck)
pALXtreme-1a‡	T7	Kanamycin (Kan)/ 100 μg mL^{-1}	ColE1 (pBR32)	Derived from pET28a(+), Δ*lacI^Q* (Blanusa et al., 2010).

Recombinant bacterial strains produced and used in this study are shown as vector maps in Figures 2.1 and 2.2.

---

† *E. coli* BL21 (DE3) lacI^{Q1} strain which contains the lacI repressor gene integrated into the genome of *E. coli* BL21 (DE3).

‡ pALXtreme vector was always used in combination with the genetically engineered *E. coli* BL21 (DE3) lacI^{Q1} strain which contains the lacI repressor gene integrated into the genome of *E. coli* BL21 (DE3).
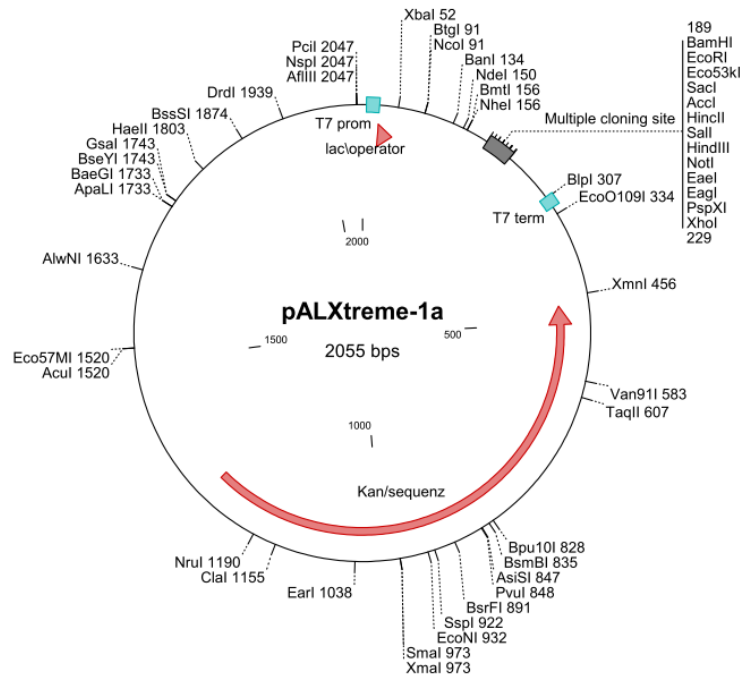
Figure 2.1. Vector map of pALXtreme-1a, a derivative of pET-28a(+). pALXtreme-1a was kindly provided by Dr. Alexander Schenk, Jacobs University, Bremen (Germany).
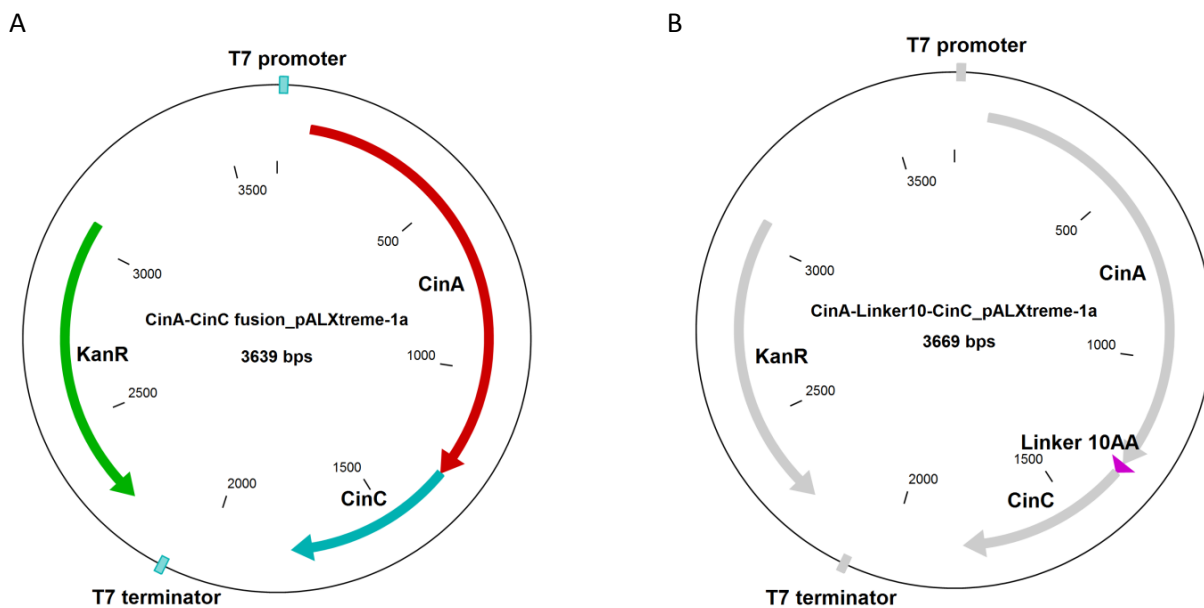
Figure 2.2. Plasmid map of CinA-CinC fusion protein and CinA-Linker 10AA-CinC fusion protein cloned in pALXtreme-1a vector. (A) C-terminus of CinA was fused to the N-terminus of CinC using PLICing cloning technique. A library of linkers from one to sixteen amino acids between the CinA-CinC fusion proteins was generated. Figure 2.2 (B) shows an exemplary plasmid map of CinA-CinC fusion protein with 10 amino acid linker (marked by magenta colored arrow).

29

2.7 Enzymes and PCR Components

Restriction enzymes, Phusion-DNA polymerase, ligase, terminal deoxynucleotidyl transferase (TdT) were purchased from New England Biolabs (Frankfurt, Germany) and were used according to supplier's recommendation. PhuS DNA polymerase and Taq DNA polymerase were prepared in-house. Catalase from bovine liver and lysozyme from chicken egg white were purchased from Sigma Aldrich Chemie (Taufkrichen, Germany).

2.8 Commercial Kits

NucleoSpin Plasmid kit	Macherey-Nagel, Düren, Germany
NucleoSpin Gel and PCR Clean-up kit	Macherey-Nagel, Düren, Germany
QIAquick PCR purification kit	Qiagen, Hilden, Germany
BCA Protein Assay kit	Thermo Fischer Scientific, Bonn, Germany
PD- 10 Desalting columns	GE Healthcare, USA

2.9 Oligonucleotides

Unmodified oligonucleotides and phosphorothioated oligonucleotides were purchased from Eurofins MWG Operon (Ebersberg, Germany) in salt-free form and dissolved in Milli-Q water to a concentration of 100 µM from which a working stock of 5 µM was prepared. Concentration of oligonucleotides used in all PCRs was 20 µM unless otherwise specified.

2.10 Software

Clone Manager 9 Professional Edition Software (Scientific & Educational Software, Cary, NC, USA) for analysis of sequencing data.

Nova 1.5 (Metrohm Autolab, USA) for electrochemical characterization of enzyme mutants.

WinCATS (CAMAG, Muttenz, Switzerland) used for semi-quantification of 2-β-hydroxy-1,8-cineole using thin layer chromatography unit.

Yet Another Scientific Artificial Reality Application (YASARA) used for generation of homology models of CinA-linker$_{(1-16)}$-CinC proteins and their variants.

3. METHODS

Molecular Biology methods were performed as described in Molecular Cloning, A laboratory Manual (Sambrook, 2001). Directed evolution methods were performed with minor modification wherever necessary according to protocols reported in Directed Evolution Library Creation: Methods and Protocols; Methods in Molecular Biology (Arnold and Georgiou, 2003).

3.1 Microbiological Methods

3.1.1 Agarose Gel Electrophoresis

DNA-fragments were separated on an agarose gel by electrophoresis in 1X TAE-buffer (TAE-buffer (50X): 2 M Tris, 1 M glacial acetic acid, 50 mM EDTA dissolved in Milli-Q H_2O pH 8.0) for 30-60 min at 90-110V depending on fragment size and gel concentration (0.8-2%). Samples from restriction endonuclease digestions, PCRs, plasmid isolation (containing approximately 25-100 ng DNA) were mixed with 1X loading dye (10 mM Tris-HCl, pH 7.6; 0.03% bromophenol blue, 0.03% xylene cyanol; 60% glycerol; 60 mM EDTA) and subsequently loaded an agarose gel (6 μL /well) along with a DNA ladder (GeneRuler 1kbp DNA Ladder, Fermentas, St. Leon-Rot, Germany) as reference. Visualization of DNA on the gel was achieved using Roti-Safe GelStain from Carl Roth (Karlsruhe, Germany) which was pre-mixed with liquid agarose before casting of the gel.

3.1.2 Polyacrylamide Gel Electrophoresis

SDS-PAGE

Protein samples were separated by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) as described previously (Sambrook et al., 1989)(Sambrook and Russell 2001). 12% SDS-PAGE were prepared composing of resolving gel (1.6 mL ddH_2O, 2 mL acrylamide mix (30%), 1.3 mL Tris (1.5 M, pH 8.8), 50 μL SDS (10%), 50 μL ammonium persulphate (APS; 10%) and 2 μL TEMED and stacking gel (1 mL) 680 μL ddH_2O, 170 μL acrylamide mix (30%), 130 μL 1.0 M Tris (pH 6.8), 10 μL 10% SDS, 10 μL 10% APS and 1 μL TEMED. Sample were prepared by incubating 10 μL of protein sample

(cell lysate or purified protein) in 6 μL SDS-PAGE loading buffer in a Thermostat (Eppendorf, Hamburg, Germany) at 95 °C for 10 min. 5 μL of sample and a pre-stained reference protein ladder (10-170 kDa, Fermentas, Germany) were loaded on the SDS-PAGE gel. The protein fractions were separated in 1 X SDS-buffer (0.3% Tris, 1.5% glycine, 0.1% SDS in H_2O) at 30 mA for 75 min using a Mini-Protean Tetra Cell system (Bio-Rad, München, Germany). For visualization of protein bands, the gels were stained in Coomassie Brilliant Blue solution (0.25% v/v coomasie, 45.5% v/v methanol, 9.2% v/v acetic acid) for 20 min followed by de-staining in 10% (v/v) acetic acid solution overnight.

3.1.3 Preparation of Chemically Competent *E. coli* cells

Chemically competent cells were prepared by inoculation of 5 mL overnight culture in LB media (37°C, 250 rpm for 16 h) with respective *E. coli* strain. Main culture of LB media (200 mL in 1 L flask) was inoculated with 1 mL preculture and incubated (37°C, 250 rpm) until an OD_{600nm} of 0.4. The culture was cooled on ice and subsequently centrifuged (4°C, 3000 rpm, 10 min; Eppendorf, Hamburg, Germany). All following steps were carried on ice and under sterile conditions. The cell pellet was resuspended in 15 mL filter sterilized ice cold TFB-1-solution (containing 30 mM potassium acetate, 50 mM manganese chloride, 100 mM rubidium chloride, 10 mM calcium chloride, 15% glycerin in Milli-Q H_2O) and incubated on ice for 10 min. The resuspended cells were centrifuged (4°C, 3000 rpm, 10 min, Eppendorf, Hamburg, Germany) and cell pellet obtained was resuspended in 2 mL filter sterilized ice cold TFB-2 solution (containing 10 mM MOPS, 75 mM calcium chloride, 10 mM rubidium chloride, 15% glycerin in Milli-Q H_2O). Competent cells were divided in 50 μL aliquots in sterile and pre-cooled Eppendorf tubes and stored at -80°C. Transformation efficiencies of up to 1×10^7 cfu μg^{-1} pUC19 were obtained.

3.1.4 Transformation of Plasmid DNA in *E. coli* cells

Transformation of DNA ligation/hybridization reaction (4 μL, around 200 ng) or plasmid DNA (20 to 30 ng) was performed using chemically competent *E. coli* cells on ice. DNA was added under sterile conditions to 50 μL frozen competent *E. coli* cells followed by

incubation on ice (10 min). DNA uptake by *E. coli* cells was achieved by heat-shock treatment (42°C for 40 s), followed by incubation on ice (2 min). Cell regeneration was achieved by addition of 900 µL SOC medium followed by incubation at 37°C for 45 min and 250 rpm. The regenerated cells were plated on LB-agar plates containing appropriate antibiotic and incubated at 37°C overnight.

3.1.5 Preparation of Cryocultures

E. coli cells harboring the desired construct were grown overnight in 5 mL culture vials (37°C, 250 rpm, 16 h). Overnight grown culture and sterile glycerol (50 % w/w in H_2O) were mixed (1:1) under sterile conditions in 2 mL Eppendorf tubes and stored at -80°C until further use. Cryocultures in microtiter plates were prepared by supplementing 100 µL/well sterile glycerol (50% w/w in H_2O) in 100 µL of overnight grown cultures (37°C, 900 rpm, 16 h, 70% humidity) and stored at -80°C until further use.

3.2 Molecular Biological Methods

3.2.1 Cloning

Two different cloning techniques were used in this thesis for sub-cloning of genes and for cloning of randomly generated mutant libraries.

Using Restriction Endonucleases

Standard restriction cloning was performed using restriction endonucleases from New England Biolabs (Frankfurt, Germany) according to manufacturer's instructions. Insert and vector DNA (1 µg each) were digested for 3 h with 1 U of corresponding restriction endonuclease in buffered reaction mixtures. The restriction endonucleases were inactivated at 55°C for 30 min. The linearized DNA fragments (gene and vector) were purified either directly using a DNA clean up kit or if necessary by gel extraction and subsequent clean up (NucleoSpin Gel and PCR Clean-up, Macherey-Nagel, Düren, Germany). In case of vector, additional incubation with 0.5 U of alkaline phosphatase was performed (37°C, 60 min) in order to reduce vector background by removal of 5'-phosphate groups.

Ligation of insert and vector (molar ratio 10:1) was performed using T4 DNA Ligase (0.5 U) in 1X T4 DNA ligase buffer (Fermentas, St. Leon-Rot Germany) at 16 °C for 16 h. The ligation reaction was subsequently transformed into chemically competent *E. coli* cells.

PLICing

PLICing (Phosphorothioate-based Ligase-Independent Gene Cloning) (Blanusa et al., 2010) was carried out as an alternative method to standard restriction endonuclease digestion in order to clone randomly generated mutant libraries.

Step 1: Amplification of target gene and vector

Vector backbone (pALXtreme-1a) and target gene (*cinA-cinC fusion*) were amplified in a PCR using oligonucleotides containing 12 complementary phosphorothioated nucleotides and a high-fidelity DNA polymerase. PCRs were performed in a final volume of 50 μL containing 1X PfuS DNA polymerase buffer, 0.2 mM dNTP mix, 20 μM each forward and reverse primer, 1 U PfuS DNA polymerase and 10 ng of template-plasmid. PCR protocol: 98°C for 1 min (1 cycle); (98°C for 30 s; 65 °C for 30 s; and 72°C for 1 m 10 s (*cinA-cinC fusion*) or 90 s (pALXtreme-1a) (20 cycles); followed by a final extension at 72°C for 5 min. Subsequently, the PCR products were analyzed by agarose gel electrophoresis and purified using the PCR cleanup kit. Parental DNA was digested by incubation with *Dpn*I endonuclease at 37°C for 16 h.

Step 2: Chemical cleavage of phosphorothioated nucleotides

Cleavage of the introduced phosphorothioated nucleotides was achieved by incubation of the amplified and purified gene and vector with an iodine cleavage solution. The cleavage reaction was performed in a PCR tube and contained 0.2 pmol insert DNA and 0.12 pmol vector DNA in 1 μL of iodine cleavage solution in a total reaction volume of 5 μL. The reaction mixture was subsequently incubated for 5 min at 70°C.

Step 3: Hybridization of vector and insert

Hybridization reaction of linearized vector and insert DNA was performed by mixing equal volumes of both cleavage reactions (typically 5 μL of vector and 5 μL of insert) and

subsequent incubation at room temperature for 5 min. Hybridization reaction (4 µL) was subsequently transformed into chemically competent *E. coli* cells (Chapter 3, Section 3.2.5). SeSaM libraries were cloned in pALXtreme-1a vector by PLICing using primers P5 - P8. The primer sequence is given in Appendix, Table 1.

3.2.2 PCR

PCRs were performed in 0.2 mL thin-walled PCR tubes from Sarstedt (Nümbrecht, Germany) employing a Mastercycler Gradient PCR-machine from Eppendorf (Hamburg, Germany). As negative control, a PCR reaction without template DNA was performed. Another PCR reaction was set without primers which served as control for *Dpn*I endonuclease digestion. This reaction was treated with *Dpn*I along with amplified products and was transformed in BL21-Gold (DE3) LacI^{Q1} to check the efficiency of *Dpn*I endonuclease to digest the parental DNA strand. Colonies obtained on the agar plates plated with control reactions were considered as background. A standard PCR (50 µL) contained: 5 U of *Taq* polymerase, 0.2 mM dNTPs, 1x *Taq* buffer, template (10-30 ng), primers (20 µM each) unless otherwise specified.

3.2.3 Colony PCR

Colony PCR was performed in order to verify cloning of target gene in vector backbone. It was performed by picking a single isolated colony from agar plate which was resuspended in 15 µL Milli-Q H$_2$O and incubated (98°C for 15 min) in order to achieve cell lysis. The PCR contained 5 µL of plasmid DNA from single colony, 5 U of *Taq* polymerase, 0.2 mM dNTPs, 1x *Taq* buffer and 5 µM each forward and reverse primer in reaction volume of 12.5 µL. Following temperature program was used: initial denaturation at 98°C (2 min), followed by 25 cycles of denaturation at 98°C (30 s), annealing (at temperature optimum for primer annealing) for 30 s and elongation at 72°C for 1 min/kb followed by a final elongation at 72°C for 5 min. The PCR products were verified by agarose gel electrophoreses.

3.2.4 SeSaM Library Generation

SeSaM library of P450$_{cin}$ fusion protein was generated based on the protocol by Mundhada et al., 2011 (Mundhada et al., 2011).

Preliminary step: Generation of ds DNA template for Step 1 and Step 3

For step 1 and step 3 template-generation, a 50 µL PCR-mixture contained: 1x Phusion GC rich Buffer; (New England Biolabs); 0.2 mM of each dNTP; 0.5 µM of each primer (P1 and P2 for Step 1 template-DNA, and P3 and P4 for Step 3 template-DNA, sequences for primer, Appendix; Table 1); 0.5 U Phusion High-Fidelity DNA polymerase (New England Biolabs, Frankfurt, Germany) and 50 ng template DNA. PCR protocol: 98°C for 30 s (1 cycle); 98°C for 10 s, 58°C for 30 s, 72°C for 50 s (15 cycles); 72°C for 5 min (1 cycle). The PCR products were digested with *Dpn*I (20 U, 37°C, 3 h) and purified (Nucleospin Extract II kit from Macherey Nagel, Düren, Germany).

Step 1: **Generation of ssDNA fragment pool with random length distribution.**

The preliminary generated CinA-Linker$_{10}$-CinC fusion Step 1 template-DNA was amplified by PCR in presence of dATPαS and dGTPαS to randomly incorporate chemically cleavable phosphorothioated bonds. PCR solutions (50 µL) contained: 1 x *Taq* Pol buffer, 0.5 µM of each primer (P1 and P4 sequences for primers, Appendix, Table 1), 5 U Taq DNA polymerase, 360 ng/kb of the SeSaM template-DNA (prepared in preliminary Step 1). According to the desired library the following nucleotide combinations were used for A and G libraries. A-libraries: 0.2 mM of each dATP, dGTP, dTTP and dCTP, and 0.06 mM dATPαS; G-libraries: 0.2 mM of each dATP, dGTP, dCTP, and dTTP and 0.05 mM dGTPαS. For both setups the same PCR protocol was performed: 94°C for 2 min (1 cycle); 94°C for 15 s, 65°C for 30 s, 72°C for 50 s (15 cycles); 72°C for 3 min (1 cycle). PCR products were purified (Nucleo Extract II kit) and eluted (80 µL NE-buffer). The cleavage of phosphorothiodiester bonds was performed in SeSaM-Taq™ buffer (1x), supplemented with iodine (10 µL, 20 mM in EtOH) and incubated (30 min, 70°C). Cleaved biotinylated DNA fragments were isolated by magnetic streptavidin beads (M-PVA SAV1, Chemagen, Baesweiler, Germany) and ssDNA was purified employing the NTC binding buffer (Nucleospin Extract II kit).

Step 2: Enzymatic addition of universal base to ssDNA fragment pool

Single-stranded DNA fragments generated in Step 1 were elongated at the 3´-OH groups by TdT (deoxynucleotidyl transferase) catalyzed incorporation of dPTP (37°C, 15 min) and terminated by heat inactivation (30 min, 75°C). Reaction mixtures (25 µL) contained: 1x TdT buffer, 0.25 mM $CoCl_2$, 1 pmol of ssDNA fragment (from Step 1b), 5 pmol dPTP and 10 U TdT. The ssDNA was purified using NucleoSpin Gel and PCR Clean-up kit using NTC buffer (Macherey Nagel, Düren, Germany).

Step 3: Full length gene synthesis

Full-length *cinA-linker10-cinC* fusion genes were generated using dPTPαS elongated DNA fragments (from step 2) as primers and step 3 dsDNA as templates (generated during preliminary step). The PCRs (50 µL) contained: 80 ng of ssDNA from step 2, 80 ng of dsDNA step 3 DNA template, 1x 3D1 polymerase buffer, 0.2 mM of each dNTPs and 1 U 3D1 polymerase. PCR protocol: 94°C for 2 min (1 cycle); 94°C for 30 sec, 52°C for 60 sec, 72°C for 2 min (30 cycles). PCR products were purified using NucleoSpin Gel and PCR Clean-up kit (Machery Nagel, Düren, Germany).

Step 4: Degenerate base replacement

Universal bases were replaced by standard nucleotides during PCR amplification. The PCR solutions (50 µL) contained: 0.5 U SeSaM Taq polymerase™, 0.5 µM of each primer (P2 and P3, Primer sequence in Appendix, Table 1), 50 ng full-length DNA from Step 3, 0.2 mM of each dNTP, 1x SeSaM Taq™ buffer. The following PCR protocol was employed: 94°C for 2 min (1 cycle); 94°C for 15 s, 60°C for 30 s, 72°C for 1 min 20 s (15 cycles); 72°C for 10 min (1 cycle). The amplified PCR product was run on 0.8 % agarose gel and gel purified with QIAquick Gel Extraction kit (Qiagen, Hilden, Germany), if non-specific amplification was observed. The purified product representing final SeSaM library was cloned by PLICing in BL-21 Gold (DE3) LacI[Q1].

3.2.5 Site directed and Site Saturation Mutagenesis

Site directed mutagenesis (SDM) of the target gene was achieved using a two-step, protocol based on the QuikChange Site-Directed Mutagenesis (SDM) System developed by Stratagene (La Jolla, CA).

In the first PCR-step, linear amplification of the target DNA with a single oligonucleotide primer is performed (forward or reverse) which avoids primer dimerization. In the second PCR-step, equal parts of amplified products from first step are mixed together followed by exponential DNA amplification. Design of primers was based on recommendations of QuikChange Site-Directed Mutagenesis manual (Agilent Technologies). Both primers were complementary to each other, with a length between 25 and 45 bases and the mutation site was flanked by at least 10 to 15 correct bases.

Site-saturation Mutagenesis (SSM) was used in order to introduce a randomize codon at a specific amino acid position in the gene of interest. The procedure employed was comparable to SDM except for the use of wobble primers, which contain instead of a single point mutation a randomized (wobble) codon. Degenerated NNK-oligonucleotides were used, where N codes for G, A, T and C, and K codes for G and T, enabling the generation of all 32 base triplets. The targeted amino acid is therefore randomized to all 20 canonical amino acids. The design of oligonucleotide primers and the mutagenic procedure were similar to site directed mutagenesis. The steps and program used for SDM and SSM are as follows:

Step I: PCR (50 µL) contained: 1x PhuS DNA polymerase buffer, 1 U PhuS DNA polymerase, 0.2 mM dNTPs, template DNA (10-30 ng) and 20 µM primers. PCR program: 98°C for 2 min (1 cycle), 98°C for 30 s, annealing at 60-65°C for 30 s min and elongation at 72°C for 1 min 20 s (3 cycles).

Step II: 25 µL of forward and 25 µL of reverse reaction were mixed together and 0.5 U of PhuS DNA polymerase was added. PCR program: 98°C for 2 min (1 cycle), 98°C for 30 s, annealing at 60-65°C for 30 s and elongation at 72°C for 1 min 20 s (15 cycles) followed by final elongation at 72°C for 5 min.

The PCR products were analyzed by agarose gel electrophoresis and parental DNA was digested by addition of 10 U of *Dpn*I endonuclease and incubation at 37°C for 16 h. It was subsequently purified using a PCR cleanup kit (Macherey Nagel, Düren, Germany) and transformed in chemically competent *E. coli* cells. The incorporation of the point mutation by SDM was verified by sequence analysis.

3.2.6 Cultivation of *E. coli* for Heterologous Protein Expression in Deep Well Plates

Individual colonies were picked from a transformation agar plate after a mutagenic procedure (SSM or SeSaM) using sterile toothpicks and were transferred into individual well of a microtiter plate (MTP) prefilled with 100 μL LB_{kana}. The MTP was cultivated in a MTP-shaker (Microtone INFORS, Heinsbach, Germany) at 37°C, 900 rpm and 70% humidity for 16 h. Overnight grown cultures were supplemented with 100 μL of 50% (w/w) glycerol and the plate was stored at -80°C as master plate. In each MTP, a negative and a positive control cultivated in triplicate served as references. Wild-type (WT) enzyme or the best enzyme variant obtained from a previous round of screening was used as the positive control. Whereas, empty vector without any cloned gene served as the negative control. During screening, the activity of mutants was analyzed and compared with respect to the activity observed in the clone containing wild-type enzyme and empty vector. Mutants showing higher activity compared to the wild-type enzyme were considered for re-screening.

3.2.7 Heterologous Expression in 96-Well Deep Well Plates

Expression of CinA-CinC Linker Library and Mutant Libraries

Master microtiter plates containing individual clones of CinA-linker$_{(1-16)}$-CinC library or clones obtained from mutant libraries were duplicated using a 96-well pin replicator to preculture flat bottom microtiter plates (Grierner Bio-one, GmbH, Frickenhausen, Germany) containing 150 μL of LB_{kan} per well. The latter were incubated in a microtiter plate shaker at 37°C for 16 h (900 rpm, 70% humidity; Multitron II, Infors GmbH, Einsbach, Germany). Expression cultures of CinA-linker$_{(1-16)}$-CinC library or mutant libraries were inoculated from overnight grown pre-cultures (6 μL) and was performed in 2 mL deep well plates (Polypropylene plates, Brand GmbH, Wertheim, Germany) containing TB media (600 μl; supplemented with 17 μM kanamycin, 0.1 mM trace element solution, 0.1 mM thiamine and 0.5 mM δ-aminolevulinic acid). Expression cultures were incubated at 37°C until an OD_{600nm} of 0.6, induced with isopropyl-thio-β-galactopyranoside (0.1 mM) and incubated in a microtiter plate shaker at 27°C for 18 h (900 rpm, 70% humidity; Multitron II, Infors GmbH,

Einsbach, Germany). Expression cultures were harvested by centrifugation at 3200 g, 4 °C for 20 min and cell pellets obtained were frozen over night at -20°C.

Expression of *E. coli* Flavodoxin Reductase (Fpr)

Master microtiter plate containing individual clones of *E. coli* Flavodoxin Reductase (Fpr) was duplicated using a 96-well pin replicator to preculture flat bottom microtiter plates (Grierner Bio-one, GmbH, Frickenhausen, Germany) containing 150 µL of LB_{kan} per well. The latter were incubated in a microtiter plate shaker at 37°C for 16 h (900 rpm, 70 % humidity; Multitron II, Infors GmbH, Einsbach, Germany). Expression culture of Fpr was inoculated from overnight grown pre-cultures (6 µL) and was performed in 2 mL deep well plates (Polypropylene plates, Brand GmbH, Wertheim, Germany) containing TB media (600 µL; supplemented with 17 µM kanamycin, 0.1 mM trace element solution, 0.1 mM sodium chloride and 0.1 mM manganese chloride). Expression cultures were incubated at 37°C until an OD_{600nm} of 0.6, induced with isopropyl-thio-β-galactopyranoside (0.1 mM) and incubated in a microtiter plate shaker at 33°C, for 4 h (900 rpm, 70% humidity; Multitron II, Infors GmbH, Einsbach, Germany). Cultures were harvested by centrifugation (3200 g, 4°C for 20 min) and cell pellets were frozen overnight (-20 °C).

3.2.8 Heterologous Protein Expression in Shake Flasks

Protein expression of CinA, CinC and CinA-CinC Fusion Protein

Preculture of CinA, CinC or CinA-CinC fusion protein was grown in sterile glass culture tube containing 4 mL of LB_{kana} medium at 37°C, 250 rpm for 16 h. Expression was performed in TB media in sterile Erlenmeyer flasks (1000 mL) containing 100 mL media. The expression media was supplemented with 17 µM kanamycin, 0.1 mM trace element solution, 0.1 mM thiamine and 0.5 mM δ-aminolevulinic acid (except for CinC) was inoculated with 1% preculture and incubated first at 37°C , 250 rpm, 70% humidity (Infors HT, Minitron, Einsbach, Germany) until an OD_{600nm} of 0.6 was reached. Expression was induced by addition of isopropyl-thio-β-galactopyranoside (0.1 mM) and incubated further for 18 h at 27 °C, 250 rpm (70% humidity, Infors HT, Minitron, Einsbach, Germany). Expression cultures were harvested by centrifugation (3200 g, 4°C for 20 min) and cell pellets were frozen overnight (-20°C).

Expression of *E. coli* Flavodoxin Reductase (Fpr)

Preculture of Fpr was grown in sterile glass culture tube containing 4 mL of LB_{kana} medium at 37°C, 250 rpm for 16 h. Expression was performed in TB media in sterile Erlenmeyer flasks (1000 mL) containing 100 mL media. The expression media was supplemented with 17 µM kanamycin, 0.1 mM trace element solution and inoculated with 1% preculture and incubated first at 37°C, 250 rpm, 70 % humidity (Infors HT, Minitron, Einsbach, Germany) until an OD_{600nm} of 0.6 was reached. Expression was induced by addition of isopropyl-thio-β-galactopyranoside (0.1 mM) and incubated further for 4 h at 33°C, 250 rpm (70% humidity, Infors HT, Minitron, Germany). Expression cultures were harvested by centrifugation (3200 g, 4°C for 20 min) and cell pellets were frozen overnight (-20°C).

3.2.9 Preparation of Crude Cell Lysates

For Screening of Mutant Libraries in Microtiter Plates

Cell pellets obtained after expression (Chapter 3, section 3.2.6) in microtiter plates were thawed at room temperature and subsequently resuspended in 200 µL of phosphate buffer (50 mM KH_xPO_4, pH 7.4) for 10 min containing 1.5 mg/mL lysozyme (≥ 40,000 U/mg) from chicken egg white (Sigma Aldrich). Cell lysis was performed by incubation in a microtiter plate shaker at 37°C for 1 h (900 rpm, 70% humidity; Multitron II, Infors GmbH, Einsbach, Germany). Clear supernatants (crude cell lysates) were obtained by centrifugation at 3200 g, 4°C for 20 min and subsequently used for activity assay.

For Purification or Activity Assay of Selected CinA-CinC Fusion Variants

Cell pellets obtained after expression (Chapter 3, section 3.2.6) in shake flasks were thawed at room temperature for 10 min and subsequently resuspended in 25 mL of appropriate buffer. Pre-lysis step was performed by sonication (40% amplitude, 30 sec x 4 cycles; SONICS Vibra cell VCX-130, Zinsser Analytics, Frankfurt, Germany) followed by cell lysis using high pressure homogenation (3 cycles; 1500 bar EmulsiFlux C3, Avestin Europe GmbH). Disrupted cells were centrifuged (3200 g, 4°C for 30 min) and the supernatant obtained was further filtered through a polyvinylidene difluoride filter (0.45 µm; Ministart

RC 25 disposable syringe filters; Sartorius, Hamburg, Germany). The cleared lysate was used for purification.

3.3 Biochemical Methods

3.3.1 Screening of P450cin Mutant Libraries in Microtiter Plates

3.3.1.1 NADPH Consumption Assay

During the development of screening system, NADPH consumption assay was performed using an equimolar mixture of CinA and CinC in a 96 well microtiter plate format using crude cell lysate (prepared as mentioned in section 3.2.8). *E. coli* Fpr was supplemented in all reaction mixtures and NADPH was used as the reduction equivalent. The final reaction mixture (50 mM KH_xPO_4, pH 7.4; 250 µL) contained 50 µL of cell lysate, 50 µl of *E. coli* Fpr lysate, 1,8-cineole (6 mM) and catalase (1500 U; Catalase from bovine liver). Immediately after supplementing NADPH, the absorption was recorded at 340 nm for 1 h at intervals of 5 min in a spectrophotometer.

3.3.1.2 Spectrophotometric Detection of Indigo Formation

In vivo Detection of Indigo Formation

Indole (final concentration: 1 mM; in DMSO 0.5% v/v) was added once 30 min after IPTG induction in *E. coli* cells harboring CinA gene carrying mutation at amino acid position 81. Following *in vivo* oxidation (24 h, 27°C, 900 rpm) cells were centrifuged (4°C, 3220 g, 30 min; 5810R, Eppendorf) and the supernatant was discarded. DMSO (300 µl) was added to wells and plates were vigorously stirred (45 min, RT, 1000 rpm) until blue color formation was visible. After short centrifugation (5 min; 4°C; 3220 g,) supernatant was transferred into polypropylene microtiter plates (Greiner Bio-One GmbH) for spectrophotometric analysis (600 to 700 nm; Tecan; Männendorf, Switzerland).

In vitro Detection of Indigo Formation

In vitro detection of indigo formation was performed in a 96- well microtiter plate format using crude cell lysate for *E. coli* cells harboring CinA with amino acid mutation at position Y81 (prepared as mentioned in section 3.2.8). *E. coli* Fpr was supplemented in all reaction mixtures and NADPH was used as reduction equivalent. The final reaction mixture (50 mM KH_xPO_4, pH 7.4; 250 µL) contained 50 µL of cell lysate, 50 µl of *E. coli* Fpr lysate, 1,8-cineole (6 mM) and catalase (1500 U; Catalase from bovine liver), 1 mM; in DMSO 0.5% v/v. Reaction was performed for 2 h at RT DMSO (300 µl) was added to wells and plates were vigorously stirred (45 min, RT, 1000 rpm) until blue color formation was visible. After short centrifugation (5 min; 4°C; 3220 g,) supernatant was transferred into polypropylene microtiter plates (Greiner Bio-One GmbH) for spectrophotometric analysis (600 to 700 nm; Tecan; Männendorf, Switzerland).

3.3.1.3 Hydroxycineole dehydrogenase (CinD) Based Detection Assay

Oxidation of 2-β-hydroxy-1,8-cineole to ketocineole was detected by measuring the rate of NADH formation. The assay was performed in a 96 well microtiter plate format using crude cell lysate (prepared as mentioned in section 3.2.8). The final reaction mixture (50 mM KH_xPO_4, pH 7.4; 250 µL) contained 30 µL of cell lysate of CinA , 30 µL of cell lysate of CinC, 30 µl of *E. coli* Fpr lysate, 1,8-cineole (6 mM) and catalase (1500 U; Catalase from bovine liver). Reaction was performed for 2 h at RT for formation of 2-β-hydroxy-1,8-cineole followed by addition of cell lysate or purified CinD (50 µL) The subsequent reaction was initiated by the addition of NAD^+ (6 mM; ε_{340} = 6200 M^{-1} cm^{-1}). The reaction (2 h; RT) was monitored by UV spectroscopy (340 nm) measuring the appearance of NADH.

3.3.1.4 Thin Layer Chromatography

Thin layer chromatography (CAMAG, Muttenz, Switzerland) was used for screening of CinA-linker(1-16)-CinC libraries, SeSaM and site saturation libraries described in this thesis. TLC was employed for semi-quantification of 2-β-hydroxy-1,8-cineole formed.

3.3.1.4.1 Activity Assay Using NADPH as Reduction Equivalent

Screening for improvement in hydroxylation activity towards 1,8-cineole for CinA-linker$_{(1\ 16)}$-CinC fusion proteins and mutant libraries (SeSaM round I) using thin layer chromatography was performed in 96-deep well plate (Eppendorf, Hamburg, Germany) format using crude cell lysate (prepared as mentioned in section 3.2.8 for semi-quantifying 2-β-hydroxy-1,8-cineole production. *E. coli* Fpr was supplemented in all reaction mixtures and NADPH was used as reduction equivalent. The final reaction mixture (50 mM KH$_x$PO$_4$, pH 7.4; 250 μL) contained 50 μL of cell lysate, 50 μl of *E. coli* Fpr lysate, 1,8-cineole (6 mM) and catalase (1500 U; Catalase from bovine liver, Sigma Aldrich). Conversions were initiated by supplementing NADPH (20 μL, 0.4 mM final concentration) and performed for 2 h at 30°C in a MTP shaker (900 rpm, Microtone INFORS, Heinsbach, Germany). Reaction products were extracted with 250 μL of ethyl acetate containing 800 μM thymol as an internal standard, dried over anhydrous NaSO$_4$ and analyzed using automated TLC (Section 3.4.4) and subsequently using Gas Chromatography (Section 3.4.3).

3.3.1.4.2 Activity Assay Using Cobalt(III)Sepulchrate and Zinc dust as an Alternative Source of Reduction Equivalent

Mediator (Cobalt(III)sepulchrate) and alaternative electron source (Zinc dust) driven activity screening and re-screening of mutant libraries using thin layer chromatography was performed in 96-deep well plate (Eppendorf, Hamburg, Germany). The reactions were performed using crude cell lysate (prepared as mentioned in section 3.2.8) and automated TLC for semi-quantifying 2-β-hydroxy-1,8-cineole production. The final reaction mixture (50 mM KH$_x$PO$_4$, pH 7.4; 250 μL) contained 5 mg zinc dust dispensed in each well of the deep well plate using a resin dispenser (Mettler- Toledo, Columbus OH, USA), 50 μl of cell lysate, 1,8-cineole (6 mM), cobalt- sepulchrate (5 mM) catalase (1500 U; Catalase from bovine liver, Sigma Aldrich). The reactions were performed for 2 h at 30°C in a MTP shaker (900 rpm, Microtone INFORS, Heinsbach, Germany). Reaction products were extracted with 250 μL of ethyl acetate containing 800 μM thymol as an internal standard. The extracted reaction products were subsequently dried over NaSO$_4$ and analyzed using automated TLC (Section 3.4.4) and subsequently using Gas Chromatography (Section 3.4.3).

3.3.2 Electrochemical Characterization of Fusion Proteins and its Variants

Electrode-driven biocatalysis was performed in an electrochemical cell with a working volume of 2 mL. Platinum sheet (A=2 cm^2) was used as working electrode (WE) and counter electrode (CE). Silver-Silver Chloride electrode served as the reference electrode (RE). The RE was inserted in a Haber-Luggin capillary (internal diameter 6 mm) filled with 0.5 M sodium sulphate. The reactions were performed in phosphate buffer (50 mM K$_X$PO$_4$, pH 7.4) at room temperature with 1 µM of CinA-linker-CinC fusion protein or its variant, 6 mM 1,8-cineole, 1500 U catalase (Catalase from bovine liver, Sigma Aldrich) and Cobalt(III)sepulchrate (5 mM). To start the reaction, working electrode was polarized to -750 mV and the reactions were performed under potentiostatic control. Reactions were performed for 10, 20, 30, 40, 50, 60 min and reaction products were extracted with 1 mL ethyl acetate containing 800 µM thymol as an internal standard. The organic phase containing the reaction products was dried over anhydrous NaSO$_4$ and analyzed by GC-FID for quantification of 2-β-hydroxy-1,8-cineole.

3.3.3 Purification of Components of P450cin and Fusion Proteins

3.3.3.1 Purification of CinA and CinA-CinC Fusion Proteins

Purification of CinA and CinA-CinC fusion protein variants was performed using ÄKTA prime plus protein purification system (GE Healthcare, München, Germany) and a preparative LC column internal diameter 15 mm, length 125 mm; KRONLAB Glass Columns, TAC 15/125 PE5-AB-2, YMC Europe GmbH, Dinslaken) by anion exchange chromatography using Toyopearl 650S-DEAE Sepharose matrix (Tosoh Biosciences, Tokyo, Japan) (bed size: 15 mm x 110 mm). Expression of CinA and CinA-CinC fusion protein variants (400 ml TB media) was performed as described in section 3.2.7. All buffer used for purification were filtered (0.45 µm filters; vacuum pump) and degassed. Cell pellets obtained after expression of CinA were resuspended in loading buffer (25 mL, 50 mM Tris-HCl, pH 7.8) and lysed using ultrasonic homogenizer (40% amplitude, 30 sec x 4 cycles, SONICS, Vibra cell, VCX-130, Zinsser Analytics, Frankfurt). A further lysis step was performed with a high pressure

homogenization (3 cycles; 1500 bar EmulsiFlux C3, Avestin Europe GmbH). Clear supernatant obtained after centrifugation (3200 g, 4°C for 30 min) was further cleared by filtration through a low protein binding filter (0.45 µm; Ministart RC 25 disposable syringe filters; Sartorius, Hamburg, Germany). The clarified and filtered supernatant (25 mL) was loaded (flow rate 1 mL/min on a pre-equilibrated (20 CVs; 50 mM Tris-HCl, pH 7.8) DEAE-ion exchange column (DEAE Sepharose 650S, bed volume: 15 mm x 125 mm, 35 µM pore size; Tosoh Biosciences LLC, Tokyo, Japan). The column was washed (20 CVs loading buffer, flow rate 3 mL/min) to remove unbound protein. Protein elution was achieved by running a NaCl gradient (50 mM Tris-HCl pH 7.8; 1 M NaCl; gradient of 0% to 40% in 200 mL volume buffer, flow rate 1.5 mL/min, fraction size 1 mL). The eluted fractions were analyzed by SDS-PAGE (Size of CinA = 45 kDa, Size of CinA-CinC fusion protein = 61 kDa; 12% SDS-PAGE gel; current 20 mA; Mini-Protean Tetra Cell, Bio-Rad Laboratories GmbH, München Germany) to check the purity of the eluted protein. Purified protein samples obtained by ion exchange chromatography were concentrated using 10K Amicon Ultra-4 filters (Darmstadt, Germany) and subsequently desalted using a PD-10 desalting column (Matrix: Sephadex™ G- 25, column bed volume: 8.3 cm, bed height: 5 mL; GE-healthcare, USA). The PD-10 column was equilibrated (5 CVs; 50 mM KH_xPO_4 buffer) and protein was eluted with elution buffer (3.5 mL; 50 mM KH_xPO_4 buffer). The desalted protein was further lyophilized for 16-18 h (Christ Alpha 1-2 LD Plus; Omnilab laborzentrum GmbH & Co. KG, Germany) and stored at -20°C.

3.3.3.2 Purification of CinC

Purification of CinC was performed by anion exchange chromatography using Toyopearl 650S-DEAE matrix (Tosoh Biosciences, Tokyo, Japan) (bed size: 15 mm x 110 mm) like described for CinA with the following modifications. Column was equilibrated in loading buffer (20 CVs; 50 mM Tris-HCl, 50 mM KCl pH 7.4, flow rate 3mL/min). Subsequently, the filtered and clarified lysate (25 ml) was loaded on the column (flow rate 1 mL/min) and unbound protein was washed out with loading buffer (20 CVs, flow rate 1.5 ml/min). CinC was eluted using salt gradient (1 M KCl, 50 mM Tris-HCl, pH 7.4, gradient from 0% to 40% in 200 ml, flow rate 1.5 mL/min). Purified CinC was concentrated using 3K Amicon Ultra-4

filters (Darmstadt, Germany) and subsequently purified by a second size exclusion chromatrography step using G-Sephadex matrix (bed volume: 15 mm x 110 mm, pore size: 100-300 µM; Sigma Aldrich, Taufkrichen, Germany) using a preparative LC column (internal diameter 15 mm, length 125 mm; KRONLAB Glass Columns, TAC 15/125 PE5-AB-2, YMC Europe GmbH, Dinslaken). The column was equilibrated (20 CVs, 50 mM Tris- HCl, pH 7.4, flow rate 3 mL/min) and pooled fractions containing CinC obtained after first purification step were loaded (flow rate 1 mL/min). CinC was eluted in buffer (50 mM Tris-HCl, pH 7.4; flow rate 1.5 mL/min, fraction size 1 ml). Purity was checked by SDS-PAGE like described above for CinA. The expected size of CinC is 16 kDa. Additionally, the purity was monitored by UV spectroscopy at 456 nm. Purified CinC was subsequently desalted in a third purification step using PD-10 desalting columns with 50 mM KH_xPO_4 buffer (Matrix: Sephadex™ G- 25- 5 mL; GE-healthcare, USA). Finally, the desalted protein was lyophilized for 16 h (Christ Alpha 1-2 LD Plus; Omnilab laborzentrum GmbH & Co. KG, Germany) and stored at -20°C.

3.3.3.3 Purification of *E. coli* Flavodoxin Reductase (Fpr)

Purification of *E. coli* Fpr was performed as previously reported with the following modification to the protocol of Hawkes et al. Fpr was purified by two anion exchange chromatography steps. In the first step, using Toyopearl 650S-DEAE matrix (Tosoh Biosciences, Tokyo, Japan, bed size: 15 mm x 110 mm) and ÄKTA prime plus protein purification system (GE Healthcare, München, Germany) on a preparative LC columns (internal diameter 15 mm, length 125 mm; KRONLAB Glass Columns, TAC 15/125 PE5-AB-2; YMC Europe GmbH, Dinslaken). Column matrix was equilibrated in loading buffer (20 CVs; 50 mM Tris-HCl; pH 7.4, flow rate 3 mL/min). The filtered and clarified lysate (25 ml) was loaded on the column (flow rate 1 ml/min) and unbound protein was washed out with loading buffer (20 CVs, flow rate 1.5 mL/min). Protein was eluted by salt gradient in elution buffer (1 M NaCl, 50 mM Tris-HCl, pH 7.4, gradient from 0% to 40% in 200 ml, flow rate 1.5 mL/min). Fractions containing Fpr were pooled and purified by a second chromatography step using hydroxyapatite column (Tosoh Biosciences LLC, Tokyo, Japan; bed volume: 15 mm x 100 mm) and LC columns (internal diameter 15 mm, length 125 mm; KRONLAB Glass Columns, TAC 15/125 PE5-AB-2, YMC Europe GmbH, Dinslaken). Column

was equilibrated in loading buffer (20 CVs; 1 mM KH_2PO_4, 100 mM NaCl; pH 7.2, flow rate 3 mL/min). Pooled fractions containing Fpr obtained after first purification step were loaded on the column (flow rate 1 mL/min)and unbound protein was washed out using loading buffer (20 CVs, flow rate 1.5 mL/min). Fpr was eluted by phosphate gradient (500 mM KH_2PO_4, pH 7.2, gradient from 0% to 30% in 200 ml, flow rate 1.5 mL/min). Purity was checked by SDS-PAGE after each chromatography step (12 % SDS-gel; 20 mA, 50 W, 75 min; Mini-Protean Tetra Cell, Bio-Rad Laboratories GmbH, München Germany). Expected size for Fpr is 29 kDa. Purified Fpr was subsequently desalted in a third purification step using PD-10 desalting column with 50 mM KH_xPO_4 buffer (Matrix: Sephadex™ G- 25- 5 mL; GE-healthcare, USA). Finally, the desalted protein was lyophilized for 18 h (Christ Alpha 1-2 LD Plus; Omnilab laborzentrum GmbH & Co. KG, Germany) and stored at -20°C.

3.4 Analytical Methods

3.4.1 Determination of P450 Concentration

Concentration of active P450 monooxygenase in purified protein samples or crude cell lysates was determined by CO-differential spectrophotometric analysis according to Omura and Sato, 1964. The reduced heme-iron has an absorption maximum at 450 nm and the concentration of active P450 (heme) can be determined spectrophotometrically. The protein (crude cell lysate, purified or lyophilized protein) was diluted in phosphate buffer (50 mM K_xPO_4, pH 7.4, 3:1 dilution). The sample was subsequently reduced using sodium dithionite (approximately 2 mg). Absorption spectrum was recorded for the protein sample in the range from 400 to 500 nm with a Varian Cary 50 spectrophotometer (Agilent Technologies, Darmstadt, Germany) as reference measurement. Subsequently, the reduced sample was gently gassed with CO for 40 sec using a lecture bottle station (Sigma Aldrich, Steinheim, Germany) with subsequent recording of the differential spectrum (400 - 500 nm). The amount of P450 was calculated from the following equation

$$C_{P450} = \frac{A_{450nm} - A_{500nm}}{\varepsilon \cdot d} \quad X\ DF$$

ε: Extinction coefficient Heme-CO-complex: 91 $mM^{-1}cm^{-1}$
d: Path length; 1cm
DF: dilution factor
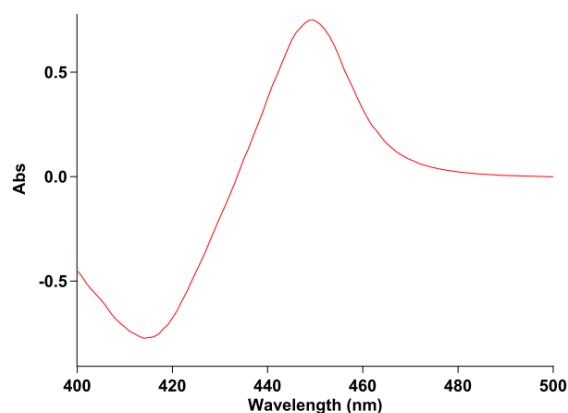
Figure 3.1.Typical CO-differential spectrum of an active P450 monooxygenase. The peak at 450 nm represents active form of P450 bound to CO monoxide after reduction with sodium dithionite.

3.4.2 Determination of FAD and FMN Concentration in CinC and *E. coli* Flavodoxin Reductase

The concentrations of CinC and *E. coli* Fpr were determined by incubating the protein samples (cell lysate or purified proteins) in a boiling water bath for 10 min at 95°C to release free FMN and FAD content. Sample was prepared by mixing the protein with phosphate buffer (50 mM K_XPO_4, pH 7.4, 1:1 dilution) and its spectrum is recorded from 400 nm to 500 nm. The sample was subsequently transferred to an Eppendorf tube and incubated in a water bath for 10 min. After cooling, the denatured protein was removed by centrifugation (20 min, 4°C, and 11,500g). The supernatant was recovered and the absorbance spectrum recorded. The concentrations of FAD and FMN were estimated by using molar extinction coefficients of 12.2 $mM^{-1}cm^{-1}$ for FMN at 446 nm and of 11.3 $mM^{-1}cm^{-1}$ for FAD at 450 nm (Aliverti et al., 1999).

3.4.3 Gas Chromatography

3.4.3.1. Extraction of Reaction Products for GC Analysis

The reaction products of P450cin conversions were extracted from the reaction mixture by two phase extraction employing ethyl acetate as extraction solvent. In order to compensate for errors due to evaporation and pipetting, a defined amount of internal standard - thymol (800 µM) was added to the extraction solvent. Reactions performed in MTPs (250 µL) were

extracted with 250 μL of ethyl acetate whereas reactions performed in Eppendorf tubes (1 mL) were extracted with 1 mL of ethyl acetate containing 800 μM of thymol as internal standard. After the completion of reactions, ethyl acetate was added to the reaction mixtures and it was mixed thoroughly (900 rpm, 15 min, RT). The organic phase was subsequently separated by centrifugation (4000 g, 10 min, RT). Next, the organic phase containing the product (2-β-hydroxy-1,8-cineole) was dried over Na_2SO_4 to remove traces of residual water. The water free organic phase was then centrifuged (4000 g, 5 min, RT) before injecting to the GC.

3.4.3.2 Quantification of Product Amount by GC-FID

Quantification of the reaction product (2-β-hydroxy-1,8-cineole) was performed using a Shimadzu GC-2010 gas chromatography coupled to a flame ionization detector (FID) (Shimadzu, Japan). 2-β-hydroxy-1, 8-cineole was detected with column FS-Supreme-5ms (5% Phenylpolysilphenylen-siloxane; 30 m x 0.25 mm x 0.25 μm film thickness; CS-Chromatographic Service, GmbH Langerwehe, Germany) using nitrogen as the carrier gas. A calibration curve was generated using authentic 2-β-hydroxy-1,8-cineole from 10 μM to 500 μM with a defined amount of internal standard (800 μM thymol). The separation of 2-β-hydroxy, 1, 8-cineole was achieved using following oven temperature program: from 80°C to 150°C with 5°C /min and hold at 300 °C for 1 min. The retention time of 2-β-hydroxy-1, 8-cineole was determined to be 9.3 min.

3.4.4 Thin layer chromatography

Extracted reaction products (40 μL) were spotted on a pre-coated Silica gel 60 F_{254} plates (20 cm x 10 cm; Merck KGaA, Darmstadt, Germany) using an automatic TLC sampler (CAMAG automated TLC sampler 4, CAMAG, Muttenz, Switzerland). A hexane/ethyl acetate (17:3) mixture was employed for TLC separation by means of an automated developing chamber (ADC2). The chromatogram was stained using a *p*-anisaldehyde/glacial acetic acid/conc. H_2SO_4 (0.5:60:0.5) mixture. 2-β-hydroxy-1,8-cineole was visualized on the plate by heating the plate with a heat gun. Densitometry evaluation of the TLC plates was carried

out at a wavelength of 634 nm (TLC scanner 3, CAMAG, Muttenz, Switzerland) and data analysis was performed with WinCATS software. Calibration curve for 2-β-hydroxy, 1, 8-cineole was generated in the range of 25 µM to 500 µM.

4. RESULTS AND DISCUSSION

The results and discussion chapter is divided in six parts. **Part I** describes the development of a screening system for determination of activity for P450cin. It focuses on various strategies investigated to develop an assay for screening of P450cin mutant libraries. **Part II** focuses on the construction of CinA and CinC fusion protein with a linker of variable length from 1 to 16 amino acids. It also describes the screening of the linker library and activity determination of the linker library towards 1,8-cineole. **Part III** gives preliminary insights describing the role of the length of the linker region in the fusion protein and its effect on the activity of the fusion protein. **Part IV** describes directed evolution of P450cin fusion protein for mediated electron transfer. The part describes random mutagenesis of P450cin fusion variant for replacement of NADPH with Zn/Cobalt (III) sepulchrate. **Part V** reports the semi-rational evolution of P450cin for mediated electron transfer. It describes redesigning of P450cin active site rationally for improvement towards Zn/Cobalt(III)sepulchrate. **Part VI** describes preliminary electrochemical characterization of P450cin variants identified from random and semi-rational evolution of P450cin. The chapter accounts the use of platinum electrode and Cobalt(III)sepulchrate for electro-biocatalytic conversion of 1,8-cineole to 2-β-hydroxy-1,8-cineole.

4.1. Development of Screening System for P450cin

4.1.1 Aim of the Work

Aim of the work described in this part of the thesis is to develop a suitable and sensitive screening system for detection of activity of P450cin towards 1,8-cineole in a 96 well microtiter plate. Screening systems employed in directed evolution experiments are usually based on the detection of substrate depletion or product formation and are typically performed on agar plates, 96 or 364 -well microtiter plates (Aharoni et al., 2005). Advanced and recently developed assay systems make use of techniques like cell surface display, emulsion technology and flow cytometry (Ruff et al., 2012). Most widely used screening systems for P450 monooxygenases are based on spectrophotometric detection of depletion of co-factor (NADH or NADPH) or are based on screening systems using surrogate substrates (e.g., pNCA) (Schwaneberg et al., 2000). However, the development of an assay for P450cin activity was challenging as no obvious change in fluorescence or absorbance is associated with hydroxycineole formation or 1,8-cineole depletion. Five different strategies were investigated for the development of a screening assay for P450cin activity towards 1,8-cineole.

4.1.2 NADPH Consumption Assay

Determination of NADPH depletion is widely employed as a screening system in directed evolution of monooxygenases. In case of P450cin, the depletion of NADPH accompanied with formation of 2-β-hydroxy-1,8-cineole can be detected spectrophotometrically at 340 nm. The reaction was performed using mixture of CinA, CinC and *E. coli* Fpr in the ratio 1:1:1 and NADPH depletion was monitored spectrophotometrically at 340 nm.

Figure 4.1.1 shows the NADPH depletion assay for P450cin using cell lysates. Activity was observed for P450cin; however the rate of reaction with a reaction time of 1 h was extremely slow. The reaction for EV and reaction without any cell lysate (negative control) did not show any depletion as expected. For the reaction without substrate there was a noticeable depletion in the absorbance of NADPH. This could be attributed to the fact that monooxygenases loose activity in absence of substrate.

The long reaction time can lead to a very high standard deviation. In addition, during the time period of 1 h, NADPH can undergo self-degradation and can lead to inaccurate measurements. Furthermore, uncoupling reaction associated with use of NADPH will further decrease the reliability for the screening system. Thus a product based detection system is favorable either using natural substrate or using a dummy substrate. As a result, the NADPH depletion assay was not investigated further.
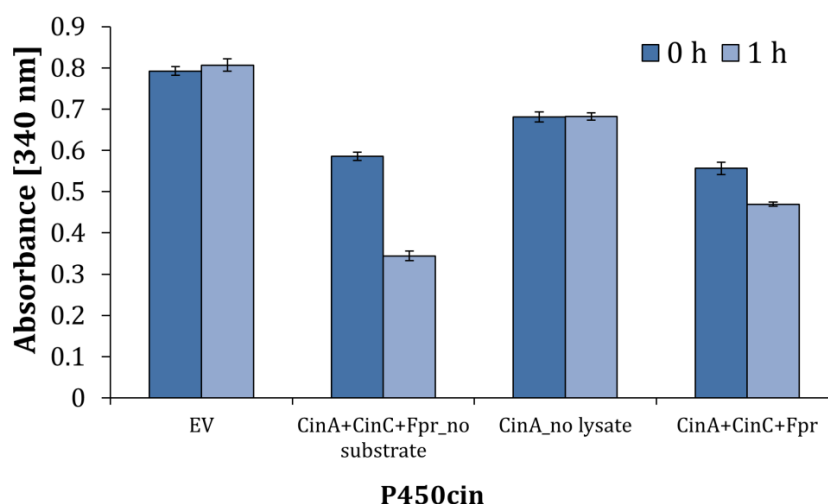
Figure 4.1.1. Spectrophotometric detection of NADPH at 340 nm using P450cin.

4.1.3 Spectrophotometric Detection of Indigo Formation

Güray Güven (Güray Güven 2010; PhD thesis, Jacobs University, Bremen) reported mutations at amino acid position Y81 in the heme domain (CinA) (Figure 4.1.2 A) resulting in mutants catalyzing the hydroxylation of indole to 3-hydroxyindole. 3-hydroxyindole undergoes spontaneous air oxidation to produce the insoluble dye indigo and indirubin (scheme 3).
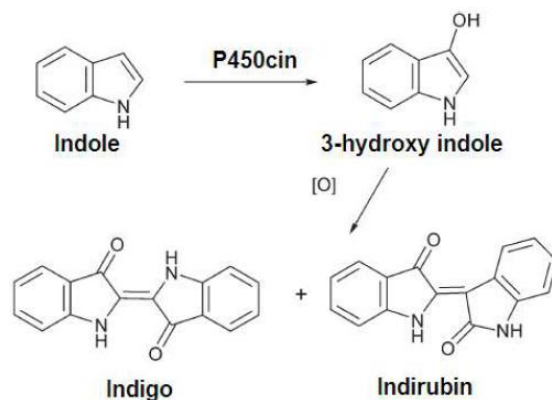
Scheme 3. Hydroxylation of indole as catalyzed by P450cin. Adapted from Celik et al (Çelik et al., 2005).
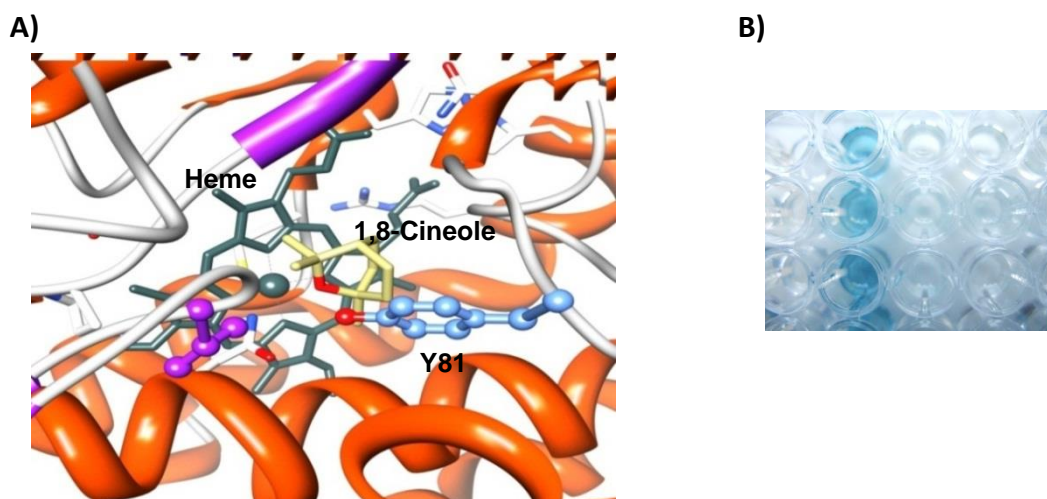
Figure 4.1.2. P450cin as indole hydroxylating catalyst. (A**)** Active site of P450cin depicting the amino acid position Y81; 1,8-cineole and heme. **(B)** Wells of a 96- well microtiter plate showing the formation of indigo dye from indole for detection at 600 nm.

The detection of indigo (formation of blue colour - Figure 4.1.2 B) can be monitored at 600 nm enabling a medium throughput screening using a microtiter plate. Indigo formation has been used in several directed evolution experiments as a pre-screening on solid phase or in microtiter plates (Çelik et al., 2005).

In order to determine the amino acids responsible for activity towards indole a SSM library was generated at position Y81. Screening of the generated SSM library using whole cells of *E. coli* harboring mutation at position Y81 was performed. *E. coli* Fpr acted as the terminal

redox partner for conversion of indole to indigo in *in vivo* reactions. *In vitro* reactions were performed by adding equimolar mixtures of cell lysates of CinA (with mutation at Y81) + CinC+ *E. coli* Fpr.
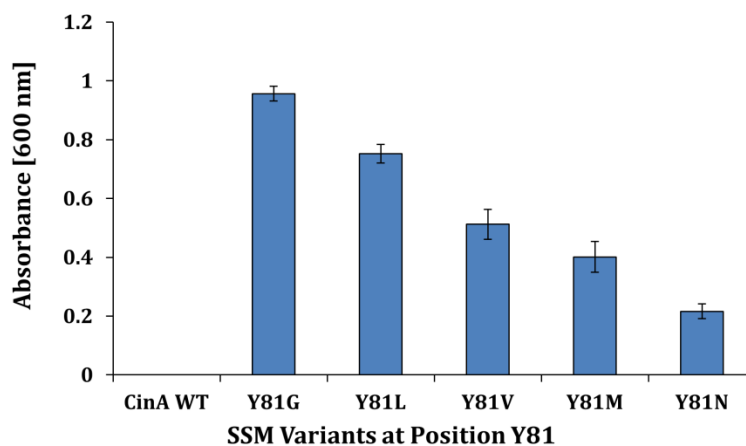
Figure 4.1.3. Activity profile for SSM variants for hydroxylation of indole to indigo using whole cells of *E. coli* harboring mutation in CinA (heme domain).

Figure 4.1.4. Activity profile for SSM variants for hydroxylation of indole to indigo using cell lysates for CinA+CinC+Fpr.

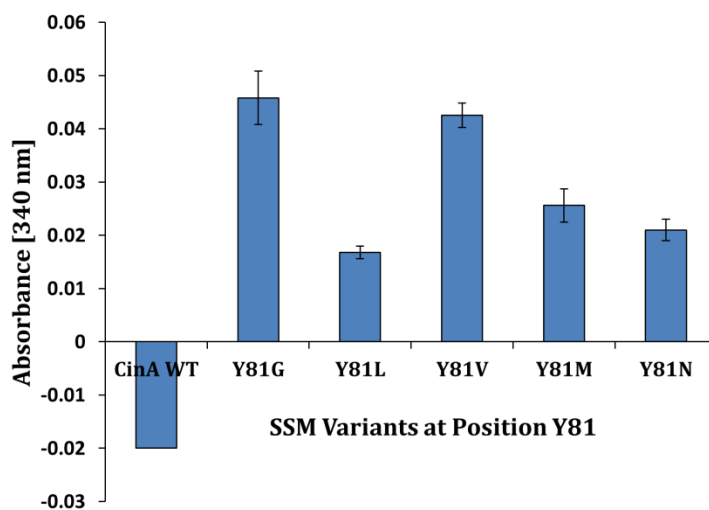Figure 4.1.3 and figure 4.1.4 show the activity profile of *in vivo* and *in vitro* reaction for conversion of indole to hydroxy indole. The *in vivo* reactions displayed the activity of SSM variants and could be detected on a 96-well plate. However, when the assay was performed

using cell lysates the conversion of indole to indigo was very low. The detection limit of the assay was not suitable for setting up a screening system with a reliable detection on a 96-well plate and hence it was not further investigated.

4.1.4 Hydroxycineole dehydrogenase (CinD) based detection assay

In the presence of an external electron transfer partner (e.g. *E. coli* Fpr) CinA and CinC catalyze the hydroxylation of 1,8-cineole to 2-β-hydroxy,1,8-cineole. The next step in 1,8-cineole degradation is believed to be the oxidation of secondary alcohol to the corresponding ketone (i.e. ketocineole) (Scheme 4; boxed).
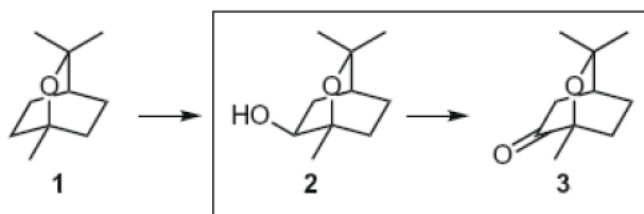
Scheme 4. Proposed pathway for 1,8-cineole **(1)** degradation to hydroxycineole **(2)** and ketocineole **(3)**. Adapted from Slessor et al (Slessor et al., 2010).

Slessor et al., demonstrated that *cinD* gene from *Citrobacter braaki* exhibited a NAD$^+$ dependent hydroxycineole dehydrogenase activity. It catalyzes the regio and enantiospecific conversion of hydroxycineole to ketocineole accompanied by the formation of NADH (Slessor et al., 2010). The latter can be detected by UV-spectroscopy and/or fluorescence measurements. By coupling the reaction of interest with ketocineole production, the levels of hydroxycineole can be indirectly monitored. Scheme 5 represents design of the screening system for detection P450cin activity indirectly by measuring the rate of formation of NADH.
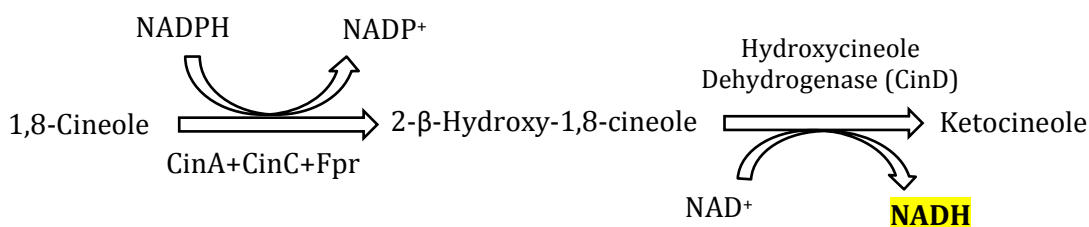
Scheme 5. Representation of CinD based screening system. 2-β-hydroxycineole-1,8-cineole produced upon oxidation of 1,8-cineole using crude cell lysate of CinA, CinC and *E. coli* Fpr was coupled with CinD and NAD$^+$ for the formation of ketocineole with concomitant release of NADH. The formation of NADH was monitored spectrophotometrically at 340 nm.

Hydroxycineole produced upon oxidation of cineole using crude cell lysate of CinA, CinC and *E. coli* Fpr was coupled with CinD and NAD$^+$ for the formation of ketocineole with concomitant release of NADH. The formation of NADH was monitored spectrophotometrically at 340 nm (Figure 4.1.5).
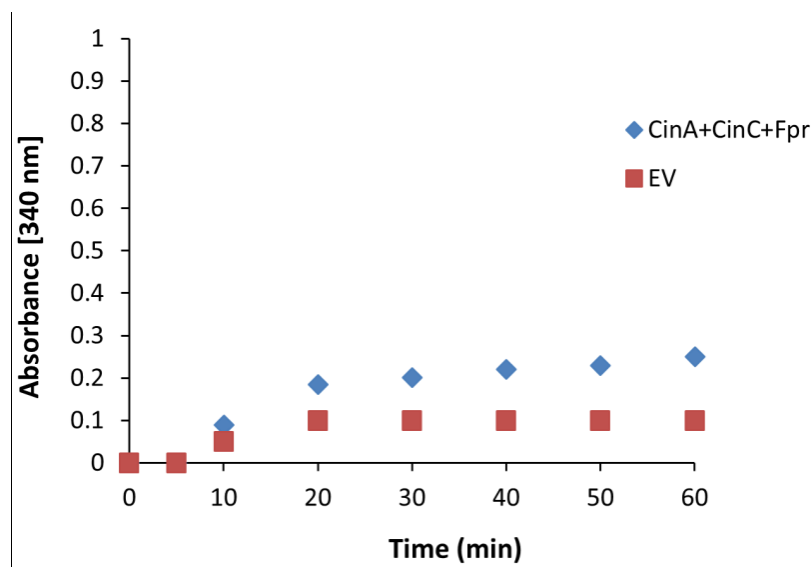
Figure 4.1.5. Spectrophotometric detection of NADH at 340 nm by coupling the oxidation of hydroxycineole to ketocineole in presence of CinD and NAD$^+$.

Figure 4.1.5 show the formation of NADH overtime. The conversion rate of hydroxycineole to ketocineole was slow (60 min). This could be also due to slower rate of conversion of cineole to hydroxycineole limiting the subsequent coupled reaction. In addition to this, the background noise from crude cell lysate of EV was very high. High background noise and

slow reaction rates made it difficult to develop a sensitive and reliable screening system on a 96-well plate and hence the CinD based assay was not pursued further.

4.1.5 Thin Layer Chromatography

The screening system using thin layer chromatography (TLC) is based on the detection of 2 β-hydroxy-1,8-cineole by separating the extracted reaction products on a TLC plate and using a mixture of p-anisaldehyde/glacial acetic acid/91 % H_2SO_4 (0.5:60:0.5) as the staining reagent. A blue colored spot corresponding to 2-β-hydroxy-1,8-cineole was observed after the TLC plate is heated with a heat gun at 100 °C for 2-3 min (Figure 4.1.6). Lane 20 shows the formation of 2-β-hydroxy-1,8-cineole using WT enzyme; Lane 1-19 shows the formation of 2-β-hydroxy-1,8-cineole using P450cin variants.
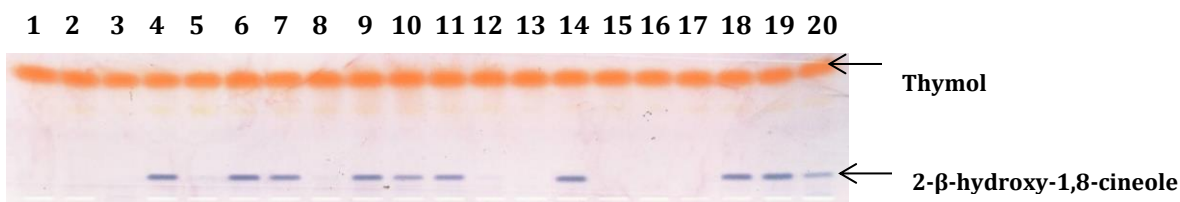
Figure 4.1.6 Illustration of a TLC plate obtained after development in mobile phase (n-hexane: ethyl acetate: 17:3) subsequent staining with p-anisaldehyde/glacial acetic acid/91 % H_2SO_4 and heating with a heat gun to develop colour for 2-β-hydroxy-1,8-cineole. Blue colored spot is 2-β-hydroxy-1,8-cineole and orange colored represents internal standard (thymol). Lane 20: formation of 2-β-hydroxy-1,8-cineole using WT enzyme; Lane 1-19: formation of 2-β-hydroxy-1,8-cineole using P450cin variants.

Validation of a screening system is an important parameter in order to determine its accuracy, efficiency and reproducibility. The proposed TLC method was validated in terms of detection range, linearity and relative standard deviation. Various parameters affect the performance of thin layer chromatography which includes material of the stationary phase, composition of the mobile phase and their interaction with the compound to be separated on the TLC. The stationary phase used for separation of 2-β-hydroxy,1,8-cineole was pre-coated silica gel 60 F254 plates (Merck® TLC plates, Merck KGaA, Darmstadt, Germany). Extracted reaction products (40 µl) were spotted using automatic TLC sampler. A mobile

phase consisting of n-hexane: ethyl acetate (17:3) was used for development of TLC plates. Using this solvent, a R_F value of 0.12 ± 0.1 were obtained for 2-β-hydroxy-1,8-cineole (Figure 4.1.7 A). Densitometric scans of the obtained bands for 2-β-hydroxy-1,8-cineole hydroxycineole was performed at the absorbance maxima at 634 nm (Figure 4.1.7 B).
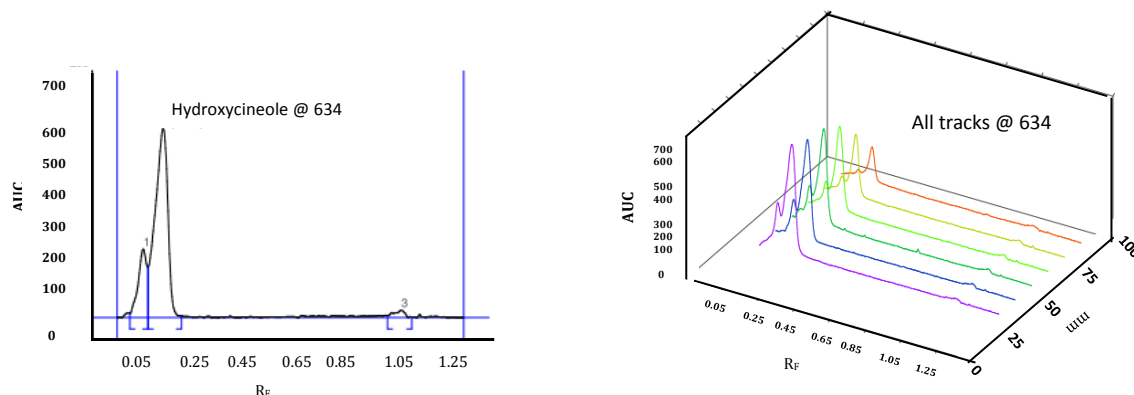
Figure 4.1.7. Analysis of 2-β-hydroxy-1,8-cineole using TLC. **A)** Graph showing the R_F of 2-β-hydroxy-1,8-cineole (peak 2) after development in mobile phase (n-hexane: ethyl acetate: 17:3). **B)** Graph displaying the peaks corresponding to 2-β-hydroxy-1,8 cineole after densitometric scan at 634 nm.

Excellent linearity was observed for calibration curve over the concentration range of 100 µM to 500 µM) and coefficient of variation was determined with respect to area and height for 2-β-hydroxy-1,8-cineole (Figure 4.1.8).

A)
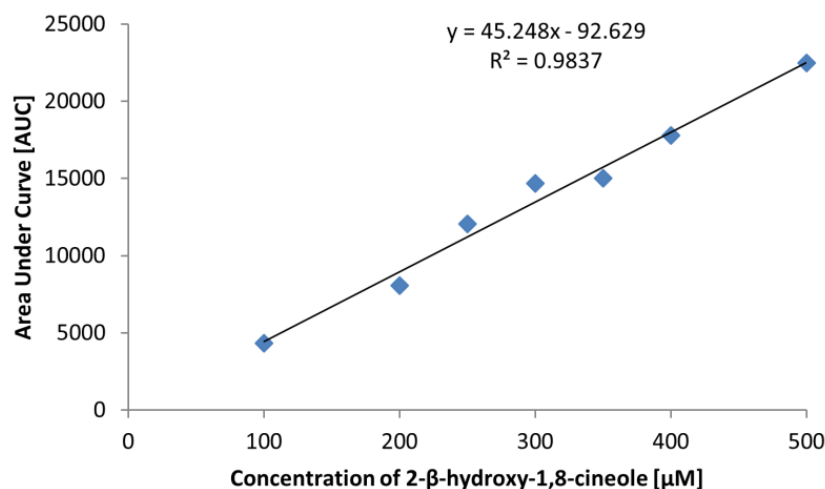
B)
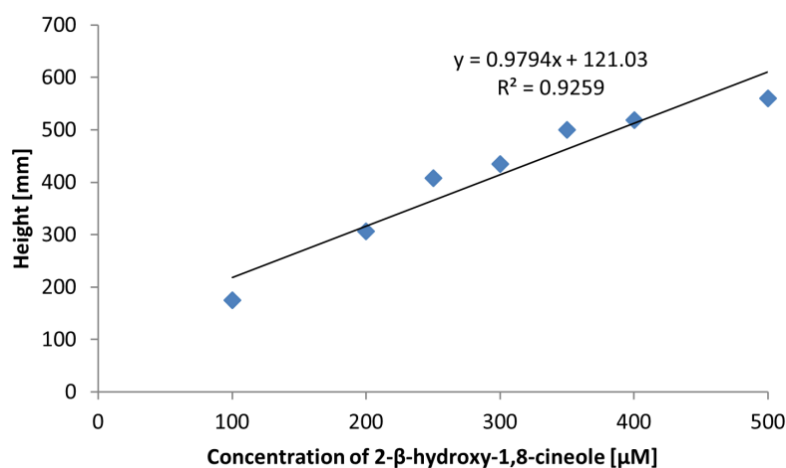
Figure 4.1.8. Calibration curve for 2-β-hydroxy-1,8-cineole using automated thin layer chromatography in a concentration range of 100 μM to 500 μM. **A)** Determined with respect to area under curve **B)** Determined with respect to height of curve.

The relative standard deviation was determined using multicomponent CinA+ CinC+ Fpr system. *In vitro* reactions were performed using NADPH containing cell lysates of CinA, CinC and Fpr for 2 h at 30 °C in all 96 wells of a 96- well microtiter plate. The extracted reaction products were spotted on the TLC plate and scanned at 634 nm. The amount of 2-β-hydroxy-1,8-cineole formed in each well of the 96-MTP was determined on the basis of relative peak area and peak height for each sample. The relative standard deviation was determined to be 18.98 % and 15.7 % for regression with area and height respectively.

The scheme for detection of 2-β-hydroxy-1,8-cineole using automated thin layer chromatography and then GC-FID is as shown in Figure 4.1.9. Though the standard deviation of the TLC based screening system was high, nonetheless it was a product based detetion system and allowed semi-quantification of 2-β-hydroxy-1,8-cineole. It allowed reliable differentiation between active and inactive clones and between WT and P450cin variants In order to quantify the amount of 2-β-hydroxy-1,8-cineole accurately, GC-FID could be employed subsequently.

CinA + CinC+ Fpr +   NADPH+ 1,8-cineole $\Longrightarrow$ 2-β-hydroxy-1,8-cineole
Or CinA-linker $_{(1-16)}$-CinC
Or Random libraries
Or SSM libraries

Extraction with ethyl acetate

Spotting of reaction product on TLC

Develpment of TLC plates

Scanning and densitometric evaluation of TLC plates for semi-quantification of 2-β-hydroxy-1,8-cineole

Quantification of 2-β-hydroxy-1,8-cineole using GC-FID
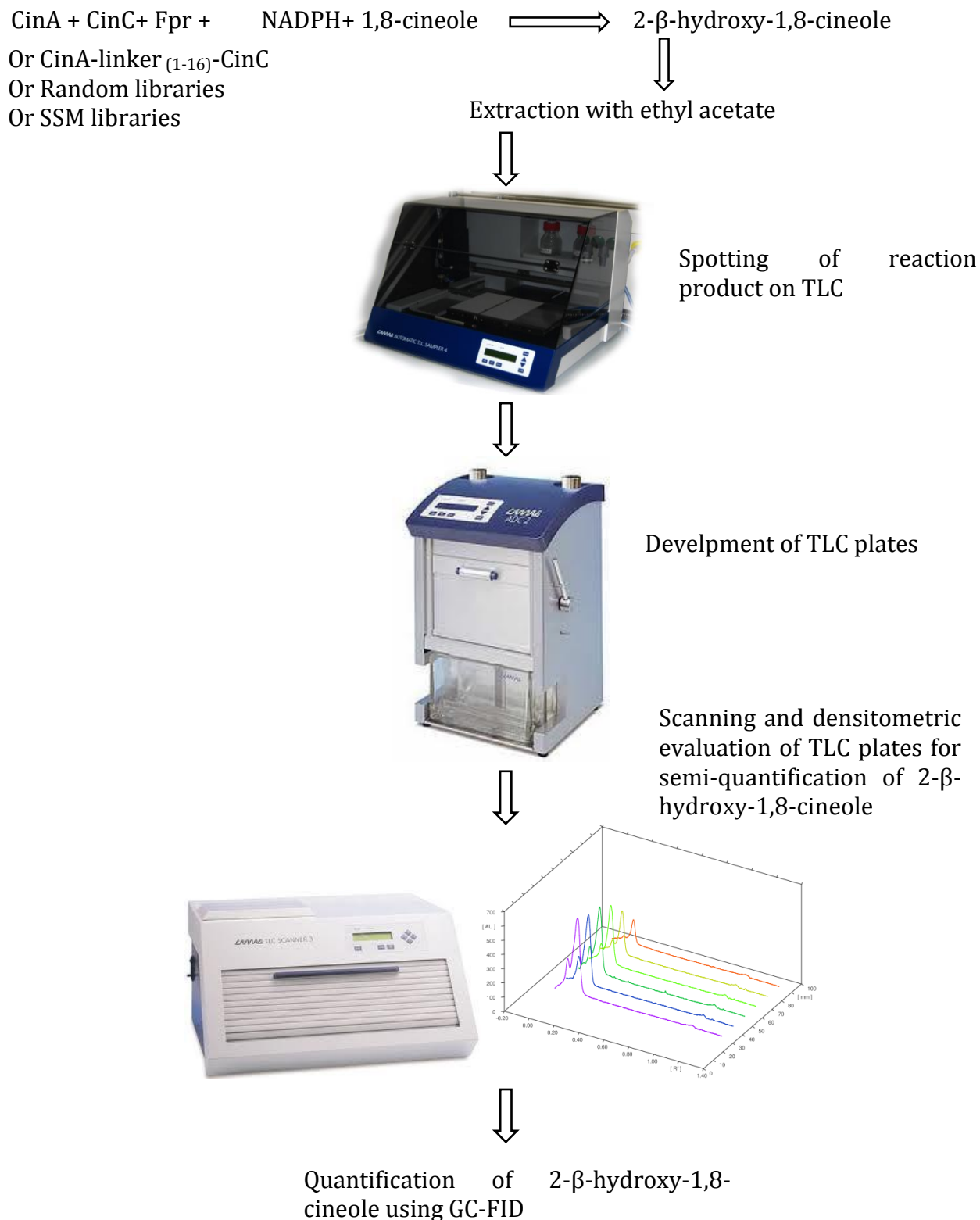
Figure 4.1.9. Steps involved in the detection of 2-β-hydroxy-1,8-cineole using automated thin layer chromatography.

4.1.6 Gas Chromatography

The TLC based screening system described above allowed semi-quantification of 2-β-hydroxy-1,8-cineole. Since it was a product based detection assay expression mutants could not be omitted and distinguished from the improved mutants. As a result, variants identified as improved with respect to the wild type using the TLC based screening system were further analyzed using GC-FID with normalized amount of P450 monooxygenase in the reaction mixture. The GC-FID measurements also allowed accurate quantification of 2-β-hydroxy-1,8-cineole. The retention time of 2-β-hydroxy-1, 8-cineole was determined to be 9.3 min.
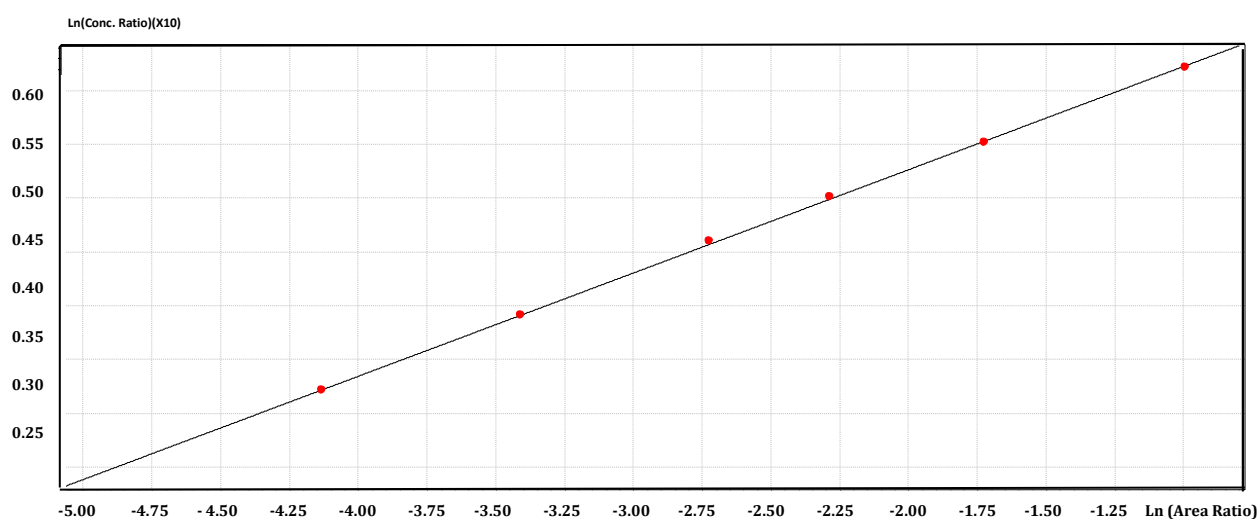
Figure 4.1.10. Calibration curve for 2-β-hydroxy,1-8-cineole using GC-FID in concentration range of 10 to 500 μM.

Figure 4.1.10 shows the calibration curve for 2-β-hydroxy-1,8-cineole in concentration range of 10 μM - 500 μM. All measurements were performed in triplicate. The relative standard deviation was determined to be 6.32 % and the co-efficient of determination (R^2) was obtained as 0.9998.

4.1.7 Discussion

One of the crucial steps in any directed evolution campaign is the development of a sensitive and reliable screening system for detection of enzymatic activity. Screening of a

random or focused mutant library using pre-screens like agar plate assay to identify active and inactive clones decreases screening efforts. It is rightly believed in directed evolution experimets ''you get what you screen for'' (You and Arnold, 1996). For mutant libraries expressed in microorganisms, screening can be performed on colonies growing on a solid substrate such as agar. Solid phase screening relies on product solubilization followed by an enzymatic reaction resulting in halo formation (zone of clearance), a fluorescent product or a strong chromogenic product (Olsen et al., 2000). Subsequently, screening of selected active clones using 96- or 384- well microtiter plate based assay formats can be performed (Aharoni et al., 2005). Advanced and recently developed assay systems make use of techniques like cell surface display, emulsion technology and flow cytometry (Ruff et al., 2012). The development of an assay for detection of P450cin activity was challenging as no obvious change in fluorescence or absorbance is associated with hydroxycineole formation or cineole depletion. Additionally, the substrate range for cytochrome P450cin is very narrow. Except for terpenoid 1,8-cineole P450cin does not accept any other substrates; not even closely related terpenoids like 1,4-cineole, α-pinene, limonene and camphor.

NADPH depletion assay is a universal assay for monooxygenases and can be monitored spectrophotometrically by measuring absorbance of NADPH at 340 nm (Glieder and Meinhold, 2003). Using the P450cin system (equimolar mixture of CinA, CinC and *E. coli* Fpr) the depletion of NADPH was monitored at 340 nm. However, the reaction was extremely slow and required about 1 h to detect a decrease in absorbance. During this long reaction time, NADPH can undergo self-degradation as well as get uncoupled forming superoxide radical, water or hydrogen peroxide leading to inaccurate measurements. Thus detection of decrease in absorbance of NADPH this assay was not a reliable and efficient option for screening of mutant libraries.

Another spectrophotometric assay which was tested was based on the detection of indigo formation of P450cin mutants at 600 nm. Several oxygenases and their mutants can form blue indigo dye after indole hydroxylation. Indigo formation can be monitored visually or spectrophotometrically at 600 nm (Çelik et al., 2005). *E. coli* colonies bearing indole hydroxylating enzymes turn blue when they grow on agar plates because indole is produced endogenously by *E. coli* (Gillam et al., 2000). Indigo detection assay is convenient and commonly used as a pre-screening on agar plates in several directed evolution

experiments to differentiate between active and inactive clones (Çelik et al., 2005). Indole can also be added to liquid cultures and the dye can be extracted for better quantification of the enzymatic activity than the agar plate assay (Gillam et al., 2000). *In vivo* reactions using P450cin harboring Y81 mutation demonstrated measurable indigo formation at 600 nm (Figure 4.1.3) whereas, when reactions using cell lysates were performed, the amount of indigo formation decreased dramatically (Figure 4.1.4). Owing to the low activity using cell lysate; this assay for detection of indigo was not suitable for screening of mutant libraries of P450cin for directed evolution.

Coupled enzyme assays have been reported for a variety of enzymes including monooxygenases, dehydrogenases, transferases (Badalassi et al., 2000; Hendricks et al., 2004; Joo et al., 1999a). The *cinD* gene from *Citrobacter braaki* exhibits a NAD$^+$ dependent hydroxycineole dehydrogenase activity catalyzing conversion of hydroxycineole to ketocineole accompanied by the formation of NADH (Slessor et al., 2010). The latter can be detected by UV-spectroscopy and/or fluorescence measurements at 340 nm. However, the sensitivity of the screening systems was low, most likely due to the slow rate of conversion of cineole to hydroxycineole thus slowing down the subsequent reporter reaction.

Detection of 2-β-hydroxy-1,8-cineole was best achieved with the thin layer chromatography system. Chromatographic based screening systems like manual TLC are not much prevalent in directed evolution campaigns. This can be attributed to the fact that they are laborious and often require a subsequent assay to be performed for product quantification. Nonetheless, the automated TLC (CAMAG, Muttenz, Switzerland) employed in this thesis could be used like general MS-methodologies with a throughput of hundreds of clones per day (Geier et al., 2013). The assay was based on derivatization of hydroxyl group in the presence of *p*-anisaldehyde, sulfuric acid and glacial acetic acid which formed a blue colored product after heating at 100 °C. The derivatized product could be detected and the amount quantified by scanning at 643 nm using a TLC scanner. The relative standard deviation of 18.98 % and 15.7 % for regression with area and height respectively were in the acceptable range for screening of random libraries generated for cytochrome P450cin. TLC was subsequently used for screening of random libraries of P450cin as described in Chapter 4, section 4.4 and 4.5.

4.1.8 Conclusion

It is very crucial in a directed evolution campaign to differentiate between WT and improved variant with the screening system. However, for the assays based on detection of depletion of NADPH at 340 nm, indigo detection at 600 nm and spectrophotometric detection of NADH at 340 nm, the difference between the wild type and the mutant enzyme was not sufficient enough for accurate quantification of enzymatic activities.

A medium throughput screening system based on detection of 2-β-hydroxy-1,8-cineole using a thin layer chromatographic detection system was established. It allowed screening of up to 200 clones / day. Since it is product detection based assay it could be applied for *in vivo* as well as *in vitro* reaction conditions. Though the throughput of the screening system was low, nonetheless it allowed accurate semi-quantification of 2-β-hydroxy-1,8-cineole with a standard deviation of was 18.98 % and 15.7 % for regression with area and height respectively. This is the first reported TLC based activity assay employed for screening of mutant libraries in a directed evolution campaign. Since the assay is based on the derivatization of hydroxyl group with *p*-anisaldehyde, sulfuric acid and acetic acid it can also be useful for discovering and isolating newer oxygenases. The TLC based screening system can also be useful to a range of other substrates (steroids, sugars, phenols, terpenes) were a colorimetric or fluorometric screening system is not available.

4.2 P-Link: A Method for Generating Multi-component Cytochrome P450 Fusions with Variable Linker Length

4.2.1 Aim of the Work

Despite the fascinating catalytic abilities, the multi-component nature of majority of cytochrome P450 system limits their applicability in synthesis, especially in cell-free applications. Generating efficient P450 enzymes by fusing the heme and redox partners with an optimized linker length is an attractive approach to overcome this challenge. Most of the currently available protocols for generation of fusion proteins are enzymatic methods with time consuming, low efficiency and complicated protocols. In this chapter a technically simple, robust and enzyme free method P-Link (P-LinK; Protein fusion with variable Linker insertion) for generation of multicomponent cytochrome P450 fusions with variable linker lengths using multicomponent cytochrome P450cin as a model enzyme system is reported. Cytochrome P450cin consists of CinA (Heme active center), CinC (FMN containing reductase) and CinB (FAD containing reductase). P-Link was used to generate a fusion protein of CinA and CinC with a variable length from 1 to 16 amino acids between the fused domains.

CinB (cindoxin reductase) is not expressible in an active form and is replaced in activity assays measurements with *E. coli* Fpr which is highly expressed in an active form in *E. coli*.

The generation of fusion protein is based on principle of phosphorothioate ligase independent cloning method, called PLICing which was previously developed by Blanusa et al (Blanusa et al., 2010). PLICing allows sequence independent cloning of DNA fragments with high efficiency and a low background of false positive colonies. PLICing employs phosphorothioated nucleotides for amplification of target gene and vector (step 1). The use of chemical cleavage step using iodine and ethanol (step 2) to generate specific complementary DNA overhangs for subsequent hybridization reaction (step 3) circumvents

the limitations of the widely employed restriction ligation cloning. Figure 4.2.1 is the schematic representation of the steps involved in PLICing cloning technique.
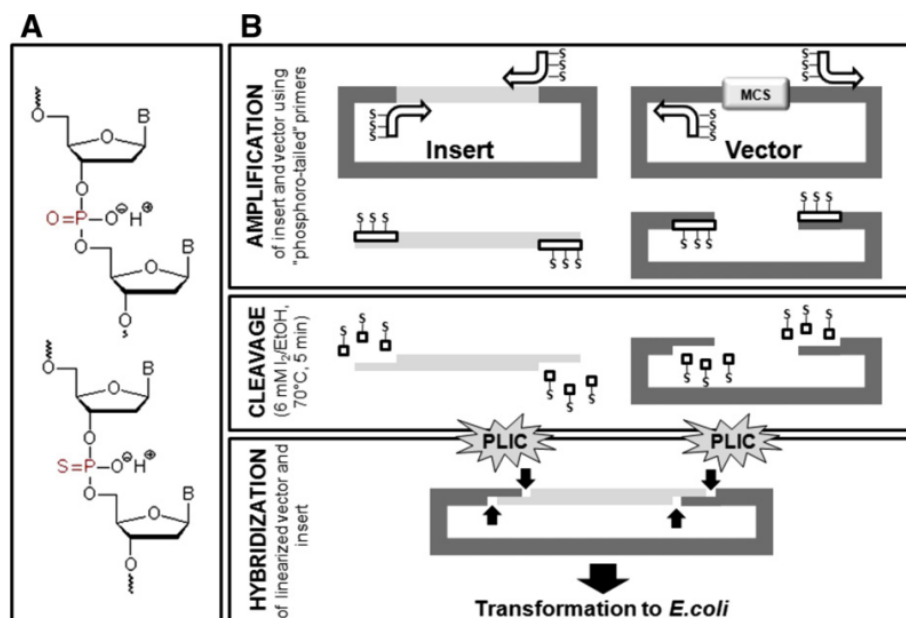
Figure 4.2.1. Schematic representation of PLICing method. **(A)** DNA backbone with a normal phosphodiester (top part) and a phosphorothiodiester bond in which one oxygen at α- phosphate is replaced by a sulfur atom (bottom part); B- protein backbone **(B)** PLICing method comprises of three steps: amplification, cleavage and hybridization (Figure taken from (Blanusa et al., 2010).

4.2.2 Generation of Fusion Protein

Generation of fusion protein is based on PLICing cloning technique. In the first step, amplification of genes (*cinA* and *cinC*) and vector (*pALXtreme-1a*) was achieved with modified phosphorothioated primers employing standard PCR protocol (step 1) and subsequently chemically cleaved in an Iodine-ethanol solution generating specific complementary overhangs (step 2). The cleaved DNA fragments were hybridized in the next step generating the CinA-CinC-pALXtreme construct (step 3). Figure 4.2.2 is the schematic representation of the steps involved in generation of CinA-CinC fusion protein.

cinA (gene size 1.2 kbp) was amplified using phosphorothioated primers *FA1* and *RA1*, whereas *cinC* (gene size 471 bp) was amplified using phosphorothioated primers *FC1* and *RC1*. Vector backbone pALXtreme-1a (vector size 2.3 kbp) was amplified separately using

phosphorothioated primers *RP1* and *FP1* (Figure 4.2.2; see Appendix for primer sequences; Table 2A)*.* The primers *FA1* and *RP1* introduce complementary phosphorothioated nucleotides at the 5´ and 3´ end of the *cinA* gene and plasmid backbone. Primers *RA1* and *FC1* introduced complementary phosphorothioated nucleotides at the 3' and 5' end of *cinA* and *cinC* respectively. Oligonucleotides *RC1* and *FP1* introduced complementary phosphorothioated nucleotides at the 3´ of *cinC* and 5´ end of vector backbone.
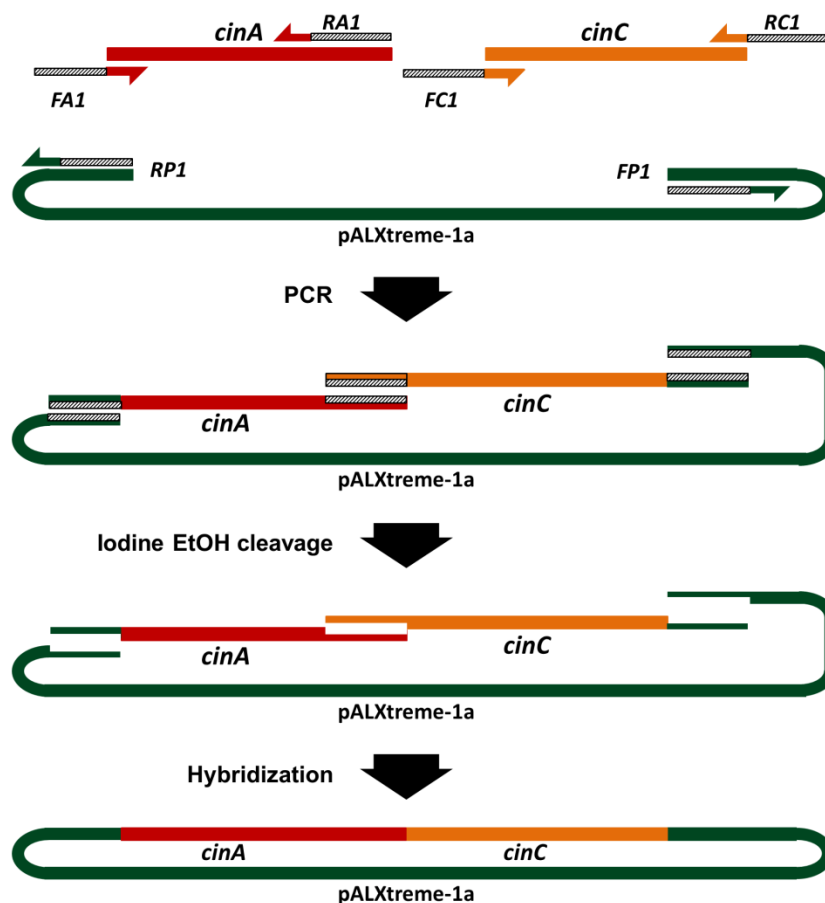
Figure 4.2.2. Steps involved in generation of the CinA-CinC fusion protein. *cinA* gene was amplified using primers *RA1* and *FA1* and *cinC* gene was amplified using primers *RC1* and *FC1.* The vector pALXtreme-1a was amplified using primers *RP1* and *FP1.* Primer *FA1*, *RP1* and primers *RC1* and *FP1* contained 12 complementary nucleotides. Subsequently the amplified products were cleaved using alkaline iodine and ethanol solution. Further hybridization of the vector and the genes generated *cinA-cinC* fusion protein.

PCR was performed for amplification of gene and vector sequence as described in chapter 3; section 3.2.2. The amplification of the gene and vector was verified on a 0.8% agarose gel. Amplified PCR products corresponding to gene size of *cinA* (1.2 kbp), *cinC* (471 bp) and vector pALXtreme-1a (2.1 kbp) were observed (Figure 4.2.3). The PCR products were digested with *DpnI* (4 h, 10 U, 37 °C) and subsequently purified using Macherey-Nagel NucleoSpin ExtractII DNA purification kit (Düren, Germany). The purified PCR products were cleaved in the presence of alkaline iodine/ethanol solution in a thermocycler as described in chapter 3; section 3.2.2 and subsequently hybridized in one PCR tube in order to generate CinA-CinC fusion protein. The hybridization mixture was transformed in chemically competent *E.coli* BL-21 Gold (DE3) lacI^{Q1} cells.
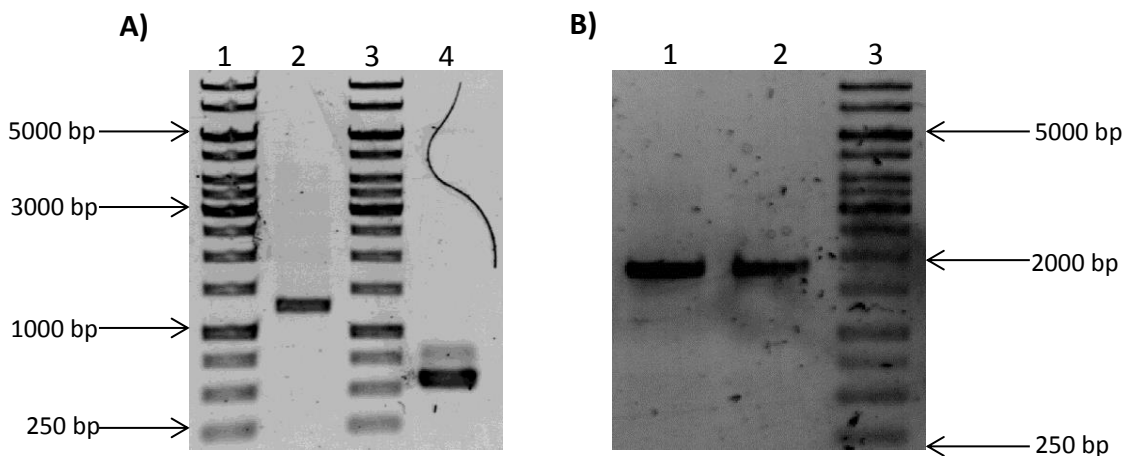
Figure 4.2.3. 0.8% agarose gel for step 1 in P-Link protocol. PCR amplification of genes (*cinA* and *cinC*) and vector (*pALXtreme-1a*). **A)** Lane 1: 1 kbp DNA ladder; lane 2: amplified *cinA* (1.2 kbp); lane 3: 10 kbp DNA ladder; lane 4: amplified *cinC* (417 bp). **B)** Lane 1 & 2: amplified pALXtreme-1a (2.1 kb); lane 3: 1 kbp DNA ladder.

4.2.3 Identification of Positive Clone for Fusion Protein

The clones obtained after the transformation of hybridized product of *cinA-cinC* genes and pALXtreme-1a vector were analyzed by colony PCR and sequencing analysis to identify a positive construct containing CinA- CinC fusion protein.

4.2.3.1 Colony PCR

Sixteen isolated colonies were picked from agar plate for colony PCR in order to determine a positive clone containing CinA-CinC fusion construct. Colony PCRs were performed as described in chapter 3, section 3.2.3 and analyzed by agarose gel electrophoreses (Chapter 3, section 3.1.3). The primers used for amplification of target gene in the fusion protein construct were primers *FA1* and *RC1* (Appendix; Table 2A for primer sequences) which amplified only the fused CinA-CinC gene (total size ~ 1.7 kb) and not the vector sequence. An expected size of ~1.7 kb was obtained in 14 out of 16 analyzed clones (Figure 4.2.4).
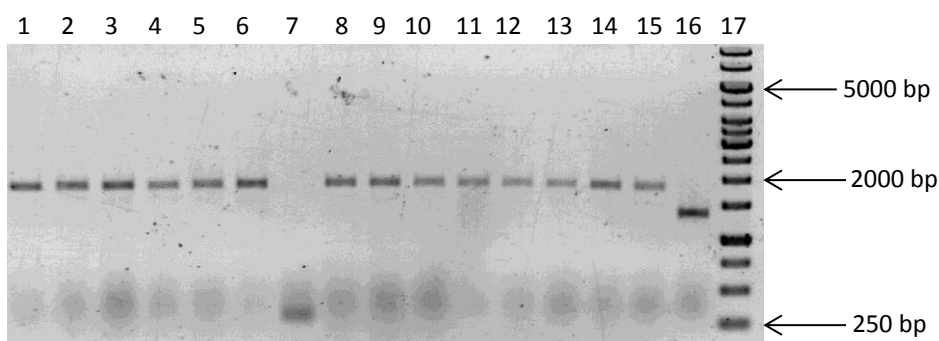
Figure 4.2.4. 0.8% agarose gel for colony PCR of fusion protein for identification of positive constructs. Lane 1-16: Colony PCR products for 16 individual colonies for identification of positive clone for CinA-CinC fusion protein; Lane 17: 1 kb DNA ladder.

4.2.3.2 Sequence Analysis of Fusion Protein

Three positive colonies identified from colony PCR for CinA-CinC fusion were sent for sequencing and obtained sequencing data was analyzed by Clone Manager software. Figure 4.2.5 shows nucleotide sequence of the CinA-CinC fusion construct.

Figure 4.2.5. Nucleotide sequence of the *cinA-cinC* fusion constructs. The black box indicates the direct joining of the nucleotide sequence in *cinA-cinC* genes. Stop codons after the *cinA* gene were removed during amplification of *cinA* gene in Step 1 of P-Link (Chapter 4; section 4.2.2).

Sequence analysis confirmed the nucleotide sequence of *cinA* was directly joined with the nucleotide sequence of *cinC* (highlighted by a black box in Figure 4.2.5) thus making a fusion of *cinA* and *cinC* genes (Chapter 4; section 4.2.2). Stop codon after *cinA* gene was removed during amplification in step 1 of P-Link in order to facilitate protein synthesis and subsequent expression of the CinA-CinC fusion protein. All three clones sequenced were positive and contained the desired CinA-CinC fusion however, only one positive clone verified by colony PCR and sequencing was used for further experiments.

4.2.4. Determination of Activity for CinA-CinC Fusion Protein

Next step after construction and expression of CinA-CinC fusion protein was determination of its activity towards 1,8-cineole. Preparation of crude cell lysate and activity assay was performed as described in chapter 3; section 3.2.9 and 3.3.1. Control reactions consisting of cell lysates of only CinA, only CinC and empty vector (1 µM each) as negative control and unfused CinA+CinC (1 µM each) as positive control in the presence of *E. coli* Fpr (1 µM) and NADPH (0.4 mM) were performed. Reactions for CinA-CinC fusion protein as well as those containing only CinA and only CinC did not show any measurable conversion of 1,8-cineole to 2-β-hydroxy,1,8-cineole. However, on the other hand, equimolar mixture of unfused CinA and CinC protein was active. The activity profiles for CinA-CinC fusion protein and control reactions are as shown in Table 4.2.1.

Table 4.2.1. Conversion of 1,8-cineole performed using CinA-CinC fusion protein supplemented with *E. coli* Fpr (1 µM) and NADPH (0.4 mM) as a reduction equivalent. Control reactions were performed using unfused CinA and CinC domains and empty vector pALXtreme-1a. (n.c. = no conversion detected).

Construct/Enzyme	Active/Inactive	Amount of 2-β-hydroxy,1,8-cineole obtained (µM)	Coupling (%)
CinA-CinC fusion	Inactive	n.c.	n.c.
CinA	Inactive	n.c.	n.c.
CinC	Inactive	n.c.	n.c.
CinA + CinC	Active	315 ± 13.88	79 ± 3.47
Empty vector	Inactive	n.c.	n.c.

The direct fusion (C terminal CinA - N terminal CinC) was inactive and conversion of 1,8-cineole to 2-β-hydroxy,1,8-cineole could not be detected. The natural fusion protein P450 BM3 has a linker sequence separating the heme and the reductase domain (Govindaraj and Poulos, 1995, 1996). Various previous studies show that separating an artificial fusion protein with a short stretch of amino acid linkers can have a considerable effect on the activity of fusion protein (Arai et al., 2001; Robin et al., 2009).

Most of the methods for insertion of linker sequences in a fusion protein are based on overlap extension PCR or using enzymes (restriction endonuclease and ligase). However they are not very efficient and only one linker length can be constructed at a time. Since the CinA-CinC fusion protein constructed was inactive; inserting linkers of different length could possibly lead to activity of the fusion protein. Thus based on PLICing cloning technique linkers of variable length from 1 to 16 amino acid residues were inserted in the CinA-CinC fusion protein.

4.2.5 Insertion of Linker Sequence in CinA-CinC Fusion Protein

For generation of CinA-CinC fusion library with a variable linker length from 1 to 16 amino acids, the plasmid containing *cinA-cinC* direct fusion construct (prepared in section 4.2.2) was used as template for subsequent PCRs (10 ng). The amino acids sequence of the linker to be inserted in CinA-CinC fusion protein was Pro-Ser-Pro-Ser-Thr-Asp-Glu-Ser-Pro-Ser-Thr-Gly-Asp-Ala-Val-Ala (PSPSTDQSPSTGDAVA). Steps involved in the insertion of linker in CinA-CinC fusion protein are amplification of cinA-cinC fusion construct (linker length 0) using phosphorothioate primers, subsequent cleavage of the phosphorothioated base pair in presence of alkaline Iodine-ethanol solution and finally hybridization (Figure 4.2.6).

The insertion of linkers in CinA-CinC fusion protein was performed in three PCRs using CinA-CinC fusion plasmid as template (Figure 4.2.6; PCR subsets i, ii, and iii).

To insert a linker length of 1 -5 amino acids (Figure 4.2.6 PCR i)*, a reverse primer (RL1)* and a mixture of forward primers *(FL1-FL5)* was employed. *RL1* includes a 34 bp sequence complementary to the 3' end of *cinA,* and contains 12 phosphorothioated nucleotides (nts) at the 5' end. *FL1* contains a nucleotide sequence complementary to the first codon of the linker, flanked by a *cinC* specific sequence at the 3' end and a phosphorothioated nt sequence complementary to the phosphorothioated sequence of *RL1* at the 5' end. Similarly,

primers *FL2*, *FL3*, *FL4, and FL5* were designed to include nucleotide sequences encoding for amino acids 2, 3, 4, and 5 of the linker. All forward primers (*FL1-FL5*) include at the 5' end a 12 phosphorothioated nts sequence complementary to the phosphorothioated sequence of the primer *RL1* (See Appendix for primer sequence, Table 2B).
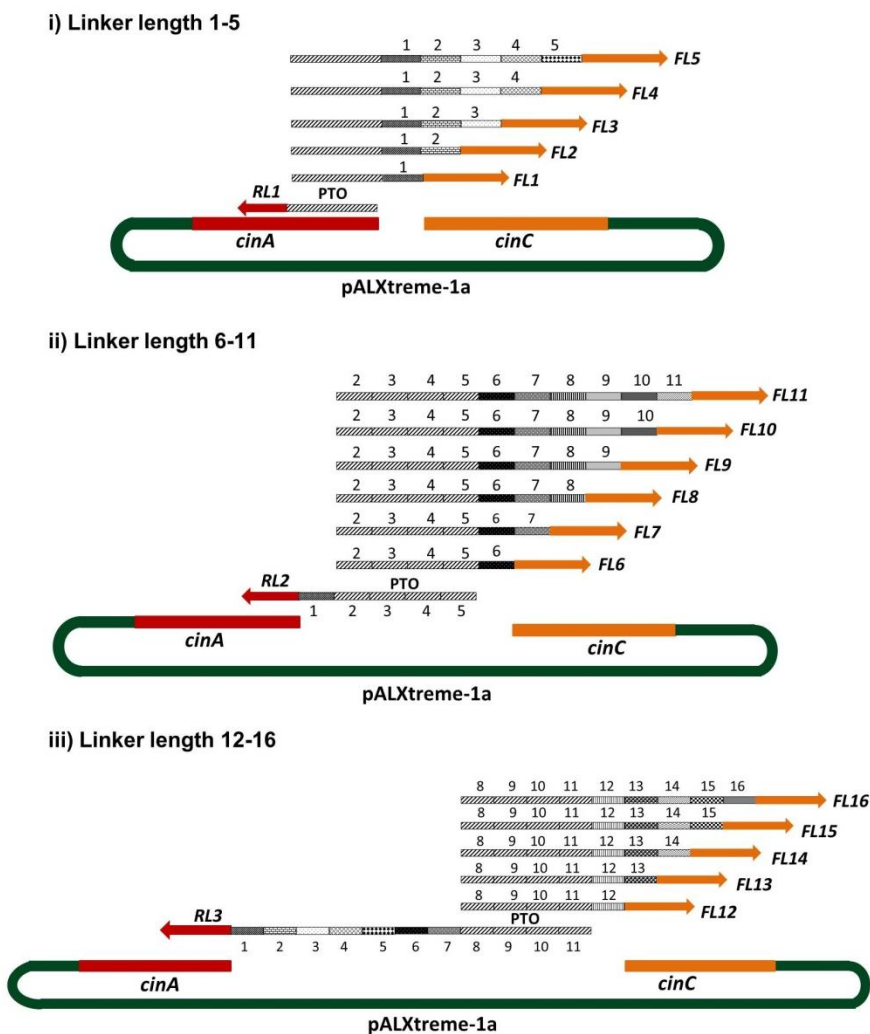
Figure 4.2.6. Steps involved in insertion of linkers of variable length between the fused CinA-CinC protein domains. Three individual PCRs were performed for insertion of linker length from 1 – 16 amino acids. PCR i performed using forward primer *RL1* and mixture of reverse primers (*FL1 – FL5*) inserted linker length between 1 - 5, PCR ii performed using forward primer *RL2* and mixture of reverse primers (*FL6 – FL11*) inserted linker length between 6 - 11 and PCR iii performed using forward primer *RL1* and mixture of reverse primers (*FL1 – FL5*) inserted linker length between 12 – 16.

In PCR ii, for the insertion of an amino acid linker length 6 - 11, the reverse primer *RL2* was designed to include a sequence coding for linker amino acid 1 - 5 (Figure 4.2.6 ii) and forward primers *FL6 - FL11* were employed using the same strategy as in PCR i to generate constructs with linker length 6 - 11.

In PCR iii, the design of the primers to insert linkers with a length of 12 - 16 amino acids employs the reverse primer *RL3*, coding for linker amino acids 1 to 10 and primers *FL12 – FL16* including the sequence coding for linker amino acids 11 to 16 (Figure 4.2.6 iii) (See Appendix for primer sequence, Table 2B). The amplification of genes from PCR i, ii and iii was verified on a 0.8% agarose gel (Figure 4.2.7). The predicted size of 3.8 kbp for the PCR product corresponding to the whole construct [gene (1.7 kbp) and vector (2.1 kbp) + linkers] was observed on the agarose gel (Figure 4.2.7).
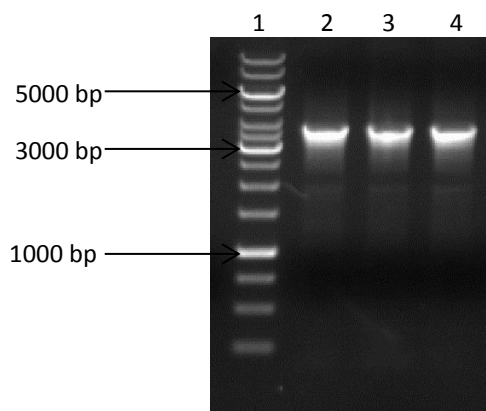
Figure 4.2.7. 0.8% agarose gel for amplification of fused *cinA-cinC* gene construct for insertion of linkers of variable length. Lane 2-4: Products from PCR i (for insertion of linker length 1 - 5-; lane 2), PCR ii (for insertion of linker length 6 - 11; lane 3) and PCR iii (for insertion of linker length 12 - 16; lane 3); Lane 1: 1 kb DNA ladder.

Amplified products obtained after PCR were digested with *DpnI* to cleave the parental DNA and were subsequently purified using Macherey-Nagel NucleoSpin ExtractII DNA purification kit for subsequent cleavage. The purified PCR fragments containing the library of *cinA-cinC* fusion with linker lengths from 1 - 5 amino acids (PCR i), linker lengths from 6 - 11 amino acids (PCR ii) and linker lengths from 12 - 16 amino acids (PCR iii) were cleaved (0.6 pmol each) in three separate PCR tubes by alkaline Iodine-ethanol solution as described in chapter 3; section 3.2.2 and subsequently hybridized in one PCR tube in order

to generate *cinA-cinC* fusion gene construct with linker lengths from 1 - 16 amino acids. Next, the hybridization mixture representing the CinA-linker$_{(1-16)}$-CinC library was transformed in chemically competent *E. coli* BL-21 Gold (DE3) lacI^{Q1} cells (Chapter 3; section 3.1.4).

4.2.5.1 Screening of the Linker Library

To determine the optimal linker length for the CinA-CinC fusion protein, the CinA-linker$_{(1-16)}$-CinC library was screened using *E. coli* Fpr and NADPH as reduction equivalent. Expression and preparation of crude cell lysate of CinA-CinC linker library was performed as described in chapter 3; section 3.2.7 and 3.2.9.

In total, 270 CinA-Linker$_{(1-16)}$-CinC variants were screened for production of 2-β-hydroxy-1,8-cineole in a 96 well microtiter plate using an automated thin layer chromatography sampler (Camag; Muttenz Switzerland). The 32 most active variants were analyzed to gas chromatography (GC-2010, Schimadzu) analysis to quantify 2-β-hydroxy-1,8-cineole formation.

The product formation of a single variant of each linker length was measured in triplicate. Linker length between 8 to 12 amino acids proved for the CinA-CinC fusion to be optimal in terms of 1,8-cineole hydroxylation. A maximum in hydroxylation activity is found at a linker length of ten amino acids (176 μM μM^{-1} P450). Interestingly a plateau is reached for variants having linker lengths of 13 - 16 amino acids (Figure 4.2.8). Fusion proteins with a linker length of 4 or less were inactive with no detectable conversion of 1,8-cineole. The activity of the CinA-$_{linker10aa}$-CinC was still lower than the activity for equimolar mixture (1 μM each) of CinA, CinC and Fpr (315 μM of 2-β-hydroxy-1,8-cineole) (Figure 4.2.8).
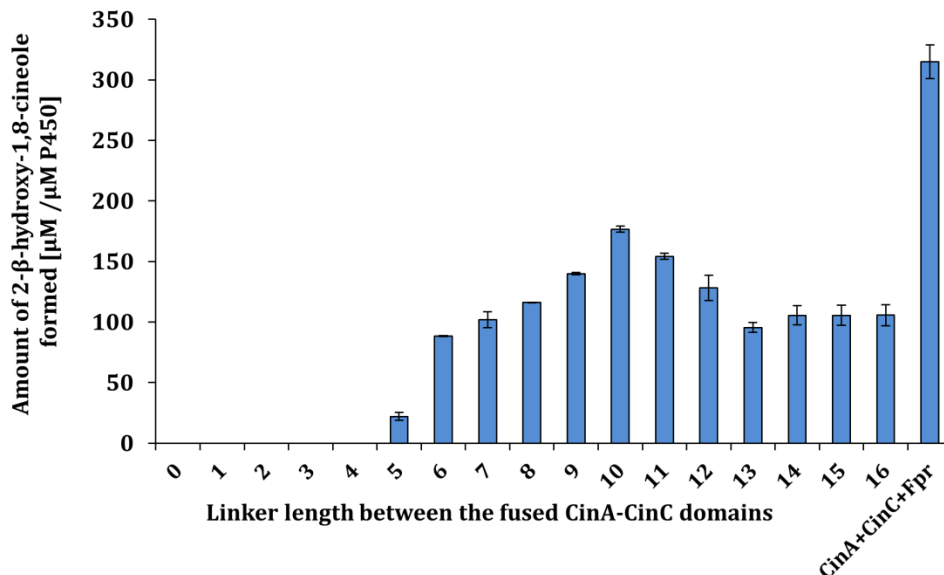
Figure 4.2.8. 2-β-hydroxy-1,8-cineole formation of CinA-CinC fusion proteins generated by P-Link *versus* linker length from 0-16 amino acids and unfused CinA+CinC+Fpr. Concentration of CinA-CinC fusion proteins are normalized to 1 µM by carbon monoxide differential scanning spectroscopy, supplemented with *E. coli* Fpr (1 µM) and reactions were initiated by addition of NADPH (0.4 mM final concentration). Highest product formation (176 µM) was reached at a linker length of 10 amino acids. In control reactions (CinA+Fpr (1 µM each), CinC+Fpr (1 µM each), empty vector+Fpr (1 µM each)) no product formation was detected. Three independent experiments were performed for variant of each linker length (1 to 16). Figure adapted from (Belsare et al., 2014).

4.2.5.2 Sequencing of Active Clones from Linker Library

Sequence analysis of the 32 most active and 15 inactive variants confirmed that all 16 possible linker lengths were generated. The CinA-CinC fusion proteins with a linker length of 4 and shorter were inactive and did not show any measurable conversion of 1,8-cineole to 2-β-hydroxy-1,8-cineole (Figure 4.2.8). Exemplary sequencing result for a fusion protein for linker length of 10 aa is shown in figure 4.2.9. Sequence analysis confirmed the nucleotide sequence of *cinA* was joined with the nucleotide sequence of linker (highlighted by a black box in Figure 4.2.9) followed by nucleotide sequence of *cinC* gene (highlighted by a blue box in Figure 4.2.9). The insertion of linker length from 1 - 16 amino acids was confirmed by sequence analysis. One clone for each linker length verified by sequencing was used for further experiments.

```
tgtgc cgattatttt tccgaaaggc aaacgtctga cgcttgcgaa ccttc tccatctact
acacg gctaataaaa aggctttccg tttgcagact gcgaacgctt ggaag aggtagatga

>.............................CinA..............>>    Linker 10

 gaccaatccc cttct atgggg taatgccctg attttatatg gcacggagac
 ctggttaggg gaaga taccc attacgggac taaaatatac cgtgcctctg

             >>...................CinC............>
```

Figure 4.2.9. Nucleotide sequence of the *cinA-cinC* fusion constructs with a linker length of ten amino acids between the CinA and Cinc protein domains. *cinA* was joined with the nucleotide sequence of linker (highlighted by a black box) followed by nucleotide sequence of *cinC* gene (highlighted by a blue box). Sequencing analysis confirmed the insertion of linker sequence in CinA-CinC fusion protein.

4.2.6 Purification of CinA-CinC Fusion Proteins, *E. coli* Fpr, CinA and CinC for Characterization

Characterization of linker variants which showed optimum product formation (CinA-linker$_{(9-11)}$-CinC was performed using purified fusion proteins. In addition, CinA, CinC and *E.coli* Fpr were also purified by anion exchange chromatography (chapter 3; section 3.3.4). Purity for all purification steps was monitored by SDS-PAGE analysis (chapter 3; section 3.1.2). Concentration of active P450 was determined by CO-differential spectroscopy (chapter 3; section 3.4.1) and concentration of active FAD and FMN was determined as described in chapter 3; section 3.4.2. Additionally, determination of total protein amount with the BCA assay was used to determine the purity of the purified protein. In all cases the purity was above 90 %.

4.2.6.1 Purification of CinA and CinA-linker $_{(9-11)}$-CinC

The purification of CinA and CinA-CinC linker variants was achieved by using anion exchange chromatography and eluting the protein with a salt gradient described in chapter 3; section 3.3. 4). Size of CinA is 45 kDa and for CinA-CinC linker variants is 61 kDa. Exemplary chromatogram of the anion exchange purification step and the SDS-PAGE analysis for purification of CinA-linker$_{10aa}$-CinC is shown in Figure 4.2.10 A.

A)

B)

C)

Figure 4.2.10. A) 12% SDS-PAGE gel for eluted fraction from anion exchange chromatographic step of CinA-linker10aa-CinC. 3 µl of eluted sample mixed with 4X loading dye was loaded in each well of the gel. Lane 1: protein loaded on column; lane 2-7: eluted fractions from the purification step; lane 8: pre-stained protein ladder from NEB. **B)** Chromatogram monitored during the purification of CinA-linker$_{10aa}$-CinC on ÄKTA by anion exchange chromatography. Blue line: UV detector signal at 280 nm; red line: conductivity; green line: concentration of elution buffer. Peak 1: impurities in the sample; peak 2: elution of CinA-CinC linker protein at 210 mM NaCl; peak 3: column washing with 1 M NaCl. **C)** 12% SDS-PAGE gel for eluted fraction from anion exchange chromatographic step of CinA. Lane 1: pre-stained protein ladder from NEB; Lane 2: protein loaded on column; lane 3: flow through; lane 4-10: eluted fractions from the purification steps.

The unbound proteins and impurities were washed with the loading buffer (Figure 4.2.10 B; peak 1) and the elution of the CinA-CinC linker protein was achieved with a concentration of 210 mM NaCl (Figure 4.2.10 A lanes 2-7 and Figure 4.2.10 B peak 2). All remaining proteins on the column were washed with 1 M NaCl (peak 3). CinA was purified in a similar manner as for CinA-linker$_{10aa}$-CinC and elution with NaCl gradient (Chapter 3; section 3.3.4; Figure 4.2.10 C).

4.2.6.2 Purification of CinC

Purification of CinC was performed as described in chapter 3; section 3.3.4. Size of CinC is 15 kDa. CinC was eluted with a salt gradient using anion exchange chromatography. The second step was a size exclusion step with isocratic elution in salt solution and the final step of buffer exchange with PD-10 desalting columns (GE healthcare Life Sciences).

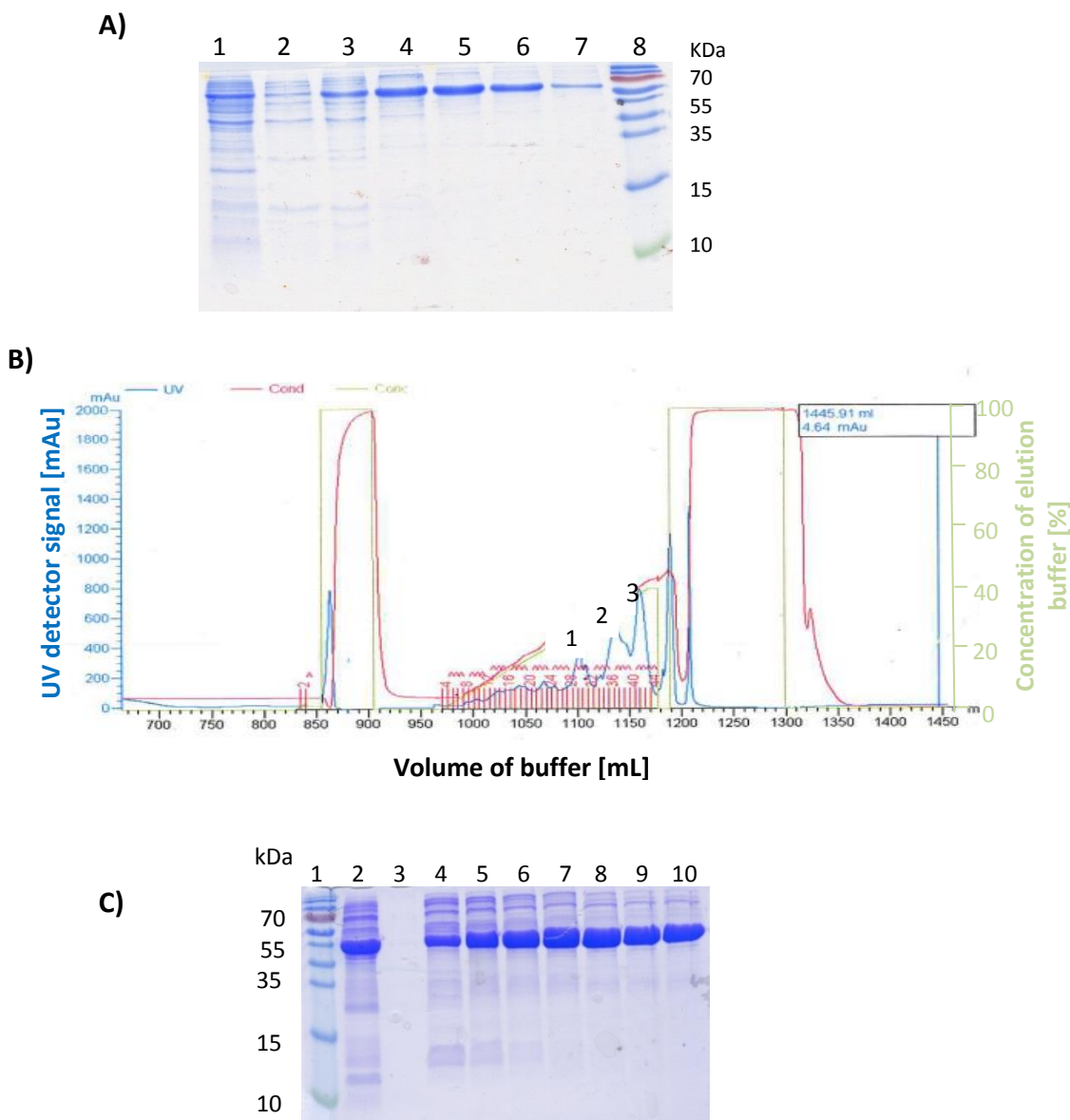**Figure 4.2.11.** 12% SDS-PAGE gel for eluted fraction from anion exchange chromatographic step CinC. 3 μl of eluted sample mixed with 4X loading dye was loaded in each well of the gel. Lane 1 - 8: eluted fractions from the purification step; lane 9: protein loaded on column; lane 10: pre-stained protein ladder from NEB.

4.2.6.3 Purification of Fpr

Purification of *E. coli* Fpr was performed as described in chapter 3; section 3.3.4. The size of Fpr is 29 kDa. Fpr was purified in two steps. The first step was an anion exchange chromatography and the protein was eluted at 240 mM NaCl. Subsequently, the protein was purified in second step using size exclusion chromatography and final step of buffer exchange with PD-10 desalting columns (GE healthcare Life Sciences).
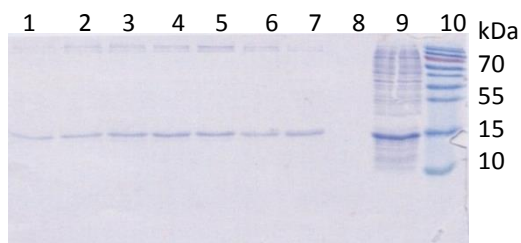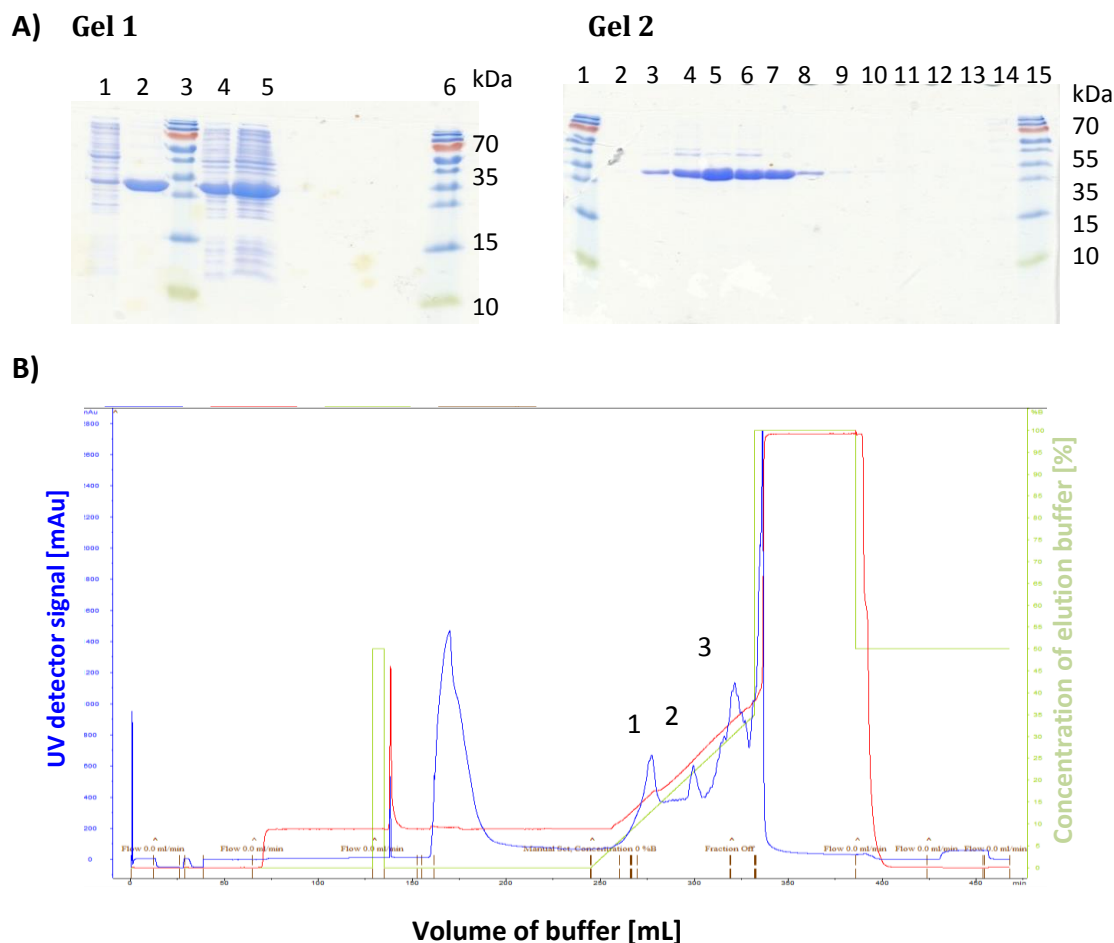
A) Gel 1   **Gel 2**

B)

Figure 4.2.12. A) 12% SDS-PAGE gel for eluted fractions from anion exchange chromatographic step of *E.coli* Fpr. 3 µl of eluted sample mixed with 4X loading dye was loaded in each well of the gel. **Gel 1**; Lane 1: unbound fraction of sample; lane 2: protein fractions from purification step 2; lane 3 & 6: pre-stained protein ladder from NEB; lane 4: protein fractions from purification step 1. **Gel 2;** Lane 1 & 15: pre-stained protein ladder from NEB; lane 2: fractions from peak 1; Lane 3-11: fractions from peak 2; lane 12-14: fractions from peak 3. **B)** Chromatogram monitored during the purification of *E. coli* Fpr on ÄKTA by size exclusion chromatography. Blue line: UV detector signal at 280 nm; red line: conductivity; green line: concentration of elution buffer. Peak 1: impurities with lower affinity for the matrix in the sample; peak 2: elution of Fpr; peak 2: impurities with higher affinity for the matrix in the sample.

4.2.7 Determination of Product Formation and Coupling Efficiency

Reactions using purified CinA-CinC fusion proteins were performed using the NADPH depletion assay. NADPH depletion associated with 2-β-hydroxy-1,8-cineole formation (1,8-

cineole depletion) was monitored by spectrophotometric measurement at 340 nm (Tecan Microtiter plate reader; Tecan Sunrise).
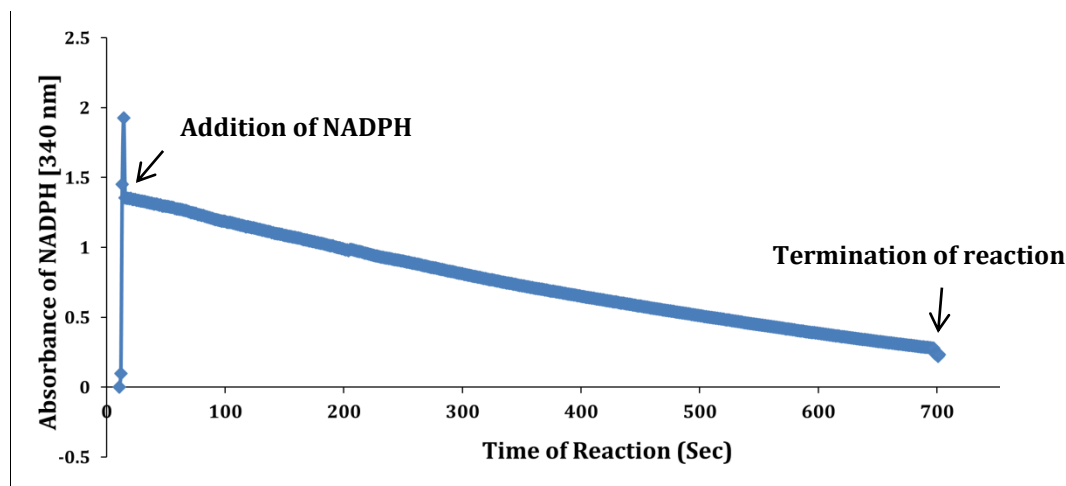

Figure 4.2.13. Depletion of NADPH monitored spectrophotometrically at 340 nm for reaction of purified CinA-CinC fusion protein with 10 amino acid linker for determination of 2-β-hydroxy-1,8-cineole formation and coupling efficiency. The graph shows the time point for addition of NADPH (initiation of reaction) and time point for termination of reaction.

The depletion (i.e. oxidation) of NADPH does not necessarily generate stoichiometric amount of product (termed coupling) but also certain reaction by-products (termed uncoupling). These reaction by-products are superoxide, water or hydrogen peroxide. High coupling (product formation) is not only desired for higher productivities but also to avoid reactive oxygen species that could potentially inactivate the enzyme. In order to determine the product formation and coupling efficiency reactions were performed using purified CinA-linker$_{(9-11)}$-CinC protein using NADPH (final concentration 0.4 mM) and purified *E. coli* Fpr. The depletion of NADPH was monitored at 340 nm and reactions were terminated at 700 sec by addition of 1 ml of ethyl acetate containing 800 μM thymol as internal standard. 2-β-hydroxy-1,8-cineole formed was determined by GC (Chapter 3; Section 3.4.3). The coupling efficiency (in percentage) was calculated as the ratio of NADPH used for substrate oxidation to the total amount of NADPH consumed by P450 in the reaction. The NADPH depletion for reaction of CinA-CinC fusion protein with 10 amino acid linker monitored at 340 nm is shown in figure 4.2.13 for determination of 2-β-hydroxy-1,8-cineole formation

and coupling efficiency. The graph shows the point where NADPH was added to the reaction and the point of termination of the reaction.
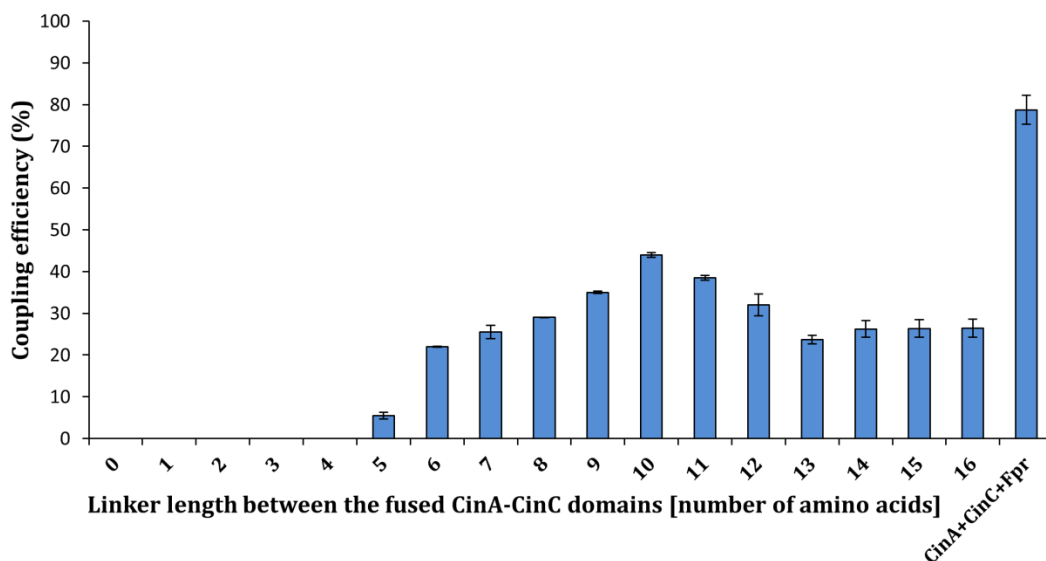
Figure 4.2.14. Coupling efficiency of CinA-linker (1-16)-CinC fusion proteins generated by P-Link *versus* linker length from 0-16 amino acids and unfused CinA+CinC+Fpr. The coupling efficiency (in percentage) was calculated as the ratio of NADPH used for substrate oxidation to the total amount of NADPH consumed by P450 in the reaction.

4.2.8 Correlation Between Thin Layer Chromatography and Gas Chromatography

TLC screening and Gas chromatography was used for screening of CinA-linker$_{(1-16)}$-CinC library as described in chapter 3, section 4.1.5 and 4.1.6. TLC screening was used as a pre-screen to differentiate between active and inactive clones using crude cell lysates.

TLC allowed quantification of amount (in μM) of 2-β-hydroxy-1,8-cineole formed in steps of 50 μM (Table 4.2.2). Additionally, TLC could quantify only 100 μM or more 2-β-hydroxy-1,8-cineole. As a result GC-FID was used to accurately quantify the amount of 2-β-hydroxy-1,8-cineole formed by each linker construct using purified enzyme. This section of the thesis describes the correlation between the two assay systems used for quantification of 2-β-hydroxy-1,8-cineole.
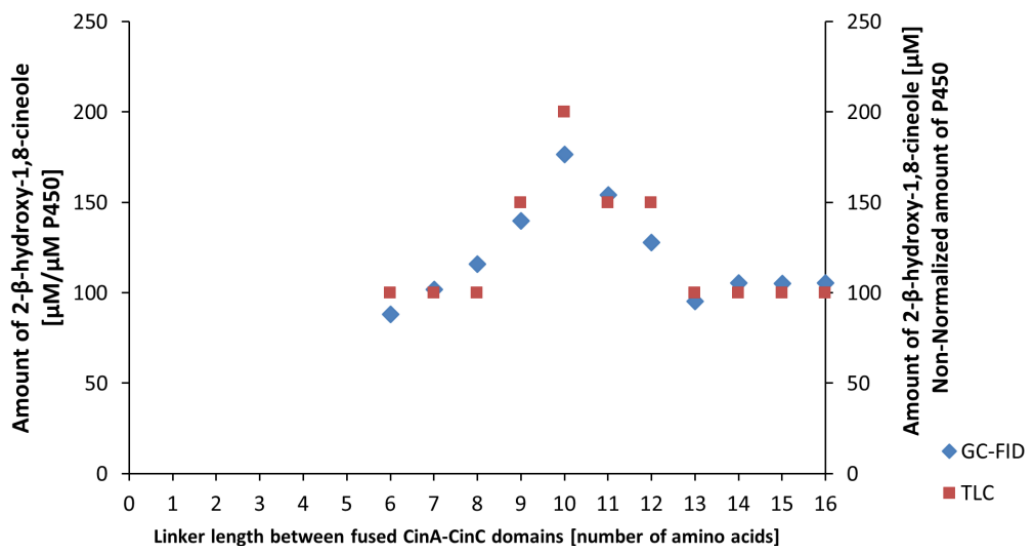
Figure 4.2.15. Correlation between measurement of amount of 2-β-hydroxy-1,8-cineole calculated by TLC and GC-FID for CinA-linker(6-16)-CinC (calculated as µM 2-β-hydroxy-1,8-cineole µM^{-1} P450).

Table 4.2.2. Amount of 2-β-hydroxy-1,8-cineole calculated by employing TLC and GC-FID . 2-β-hydroxy-1,8-cineole formed is calculated as µM µM^{-1} P450 in a reaction time of 2 h.

Linker length	Amount of 2-β-hydroxy-1,8-cineole formed [µM µM^{-1} P450] in 2 h	
	GC-FID	**TLC**
6	88.34	100.00
7	102.03	100.00
8	116.12	100.00
9	140.00	150.00
10	176.67	200.00
11	154.26	150.00
12	128.17	150.00
13	95.54	100.00
14	105.61	100.00
15	105.49	100.00
16	105.64	100.00

Screening using TLC was performed in a 96- well microtiter plate without normalizing the amount of P450 in a reaction whereas in case of purified enzymes, reactions were performed using normalized amounts of P450 (1 µM). Table 4.2.2 shows the amount of 2-β-

hydroxy-1,8-cineole as calculated by TLC and GC-FID. Quantification using TLC assay could determine the amount of 2-β-hydroxy-1,8-cineole only in steps of 50 μM and therefore it was necessary to use GC-FID for accurate quantification of 2-β-hydroxy-1,8-cineole. As seen in figure 2, a good co-relation was observed while detection of 2-β-hydroxy-1, 8-cineole using the TLC and GC-FID.

4.2.8 Discussion

Generation of fusion proteins is an attractive approach to simplify the multicomponent system (P450 part and reductase partner) of cytochrome P450s on genomic level. In addition, it is also an excellent technique to explore the catalytic potential of P450s in which the natural redox partners are still unknown. Generating fusion proteins cannot only ease handling of the system, facilitate expression studies but also simplify the system on a genetic level which can be used in directed evolution campaigns. Most of the currently used methods for generating fusion proteins are enzymatic which are time consuming, tedious and have low efficiency. In this section of the thesis P-Link a robust, technically simple and easy method for generation of fusion proteins is developed and described. P-Link employs amplification of gene using (step 1) primers containing phosphorothioated nucleotides for fusion generation and insertion of linker of variable length. The use of chemical cleavage step using iodine and ethanol (step 2) to generate specific complementary DNA overhangs for subsequent hybridization reaction (step 3).

The critical step in P-Link is the cleavage step to generate single stranded DNA overhangs using Iodine-ethanol to ensure efficient fusion generation. The chemical cleavage with Iodine-ethanol under alkaline conditions (pH 9.0) specifically cleaves PTO nucleotides. Eckstein and Gish have reported the cleavage mechanism of phosphorothioated nucleotides in DNA and RNA and its application in DNA sequencing (Eckstein and Gish, 1989; Gish and Eckstein, 1988). The proposed cleavage mechanism for phosphorothioated nucleotides through alkylation is highly regioselective and does not lead to cleavage or modification of natural phosphodiester bonds (Figure 4.2.15).
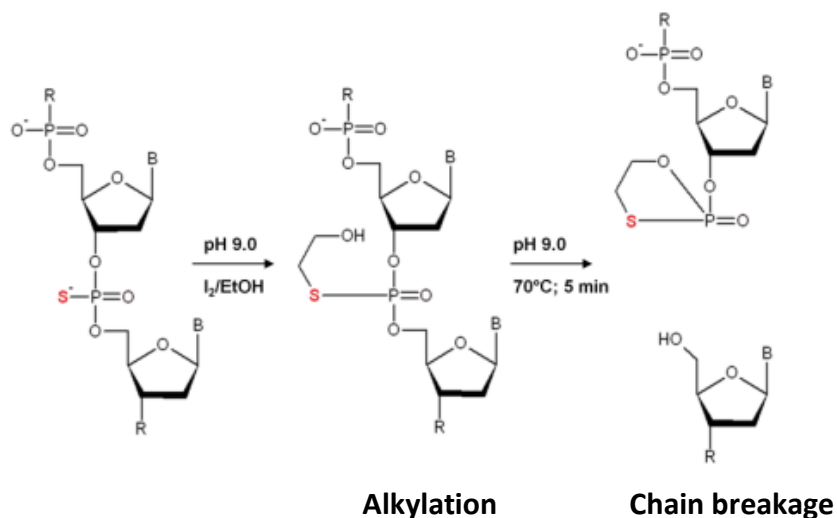
Figure 4.2.16. Mechanism of chemical cleavage of phosphorothioated nucleotides in presence of Iodine and ethanol under alkaline conditions based on the mechanism given by Eckstein and Gish 1989 and Gish and Eckstein 1988.

Purification step can be omitted prior to transformation since Iodine and ethanol do not interfere with hybridization of complementary single stranded DNA (Blanusa et al., 2010). Various methods previously reported for insertion of linkers make use of enzymes (restriction endonuclease and ligase) (Robin et al., 2009; Sibbesen et al., 1996) or based on overlap extension PCR (Urban et al., 1997) in contrast to P-Link which is based on chemical cleavage of phosphorothioate bonds of PCR products amplified using phosphorothioated primers.

P-Link is simple (no enzymatic digestion and ligation step as compared to enzymatic approached) and less laborious (16 PCRs and 16 transformation reactions for overlap extension PCR in comparison to 3 PCRs and 3 transformation for P-Link).

An alternative to P-Link would be using methods based on overlap extension PCR, P-Link can however very likely be expanded to vary linker sequence or to include different vector backbones in a single tube experiment, thereby exploring more alternatives for expression of fusion constructs. An additional advantage of P-Link is the expansion of P-Link to fuse three and four multi-component systems at once which are very likely much less efficient in case of overlap extension PCR.

The first artificial fusion protein for bacterial P450 monooxygenases was described for cytochrome P450cam (Sibbesen et al., 1996) which oxidize camphor to 5-*exo-*

hydroxycamphor. The fusion protein was constructed between P450cam, putidaredoxin (Pd) and putidaredoxin reductase (PdR), the natural reductase partners of P450cam. Several fusion proteins were constructed in which the order of three fusing domains and the linkers between them were varied. The highest product formation rate of 30 µM min^{-1} µM fusion protein^{-1} was obtained for fusion construct Pdr-Pd-P450cam separated by linkers with amino acid composition TDGASSS and PLEL respectively.

One important parameter that governs the activity of fusion proteins is the orientation of the fused protein domains and the length of the linker separating these domains (Robin et al., 2009; Sibbesen et al., 1996). In case of P450 BM3, naturally fused cytochrome P450, the C-terminal of heme domain is fused to the N-terminal of the reductase domain via short stretch of amino acid residues. The role of the linker region is not clear however, it is commonly postulated that the linker region acts as a flexible hinge allowing recognition and interaction between the two domains (Hannemann et al., 2007; Sevrioukova et al., 1999).

Furthermore, Munro et al., reported that the linker region might be also responsible for stabilization of the active dimeric form of P450 BM3 (Munro et al., 2007). Previous reports establish that the length of the linker is decisive for activity of fusion proteins rather than the amino acid composition of the linker. In two individual studies Govindraj and Poulos reported that changing Arg-Lys-Lys stretch in the middle of the linker sequence in P450 BM3 to Ala-Ala-Ala does not alter the functional properties of either of the two domains. However, deleting three or six amino acids leads to drastic effect on the hydroxylation activity of the enzyme. The three amino acid deletion mutant exhibited 10% of the hydroxylation activity while the six amino acid deletion mutant showed nearly undetectable levels of fatty acid hydroxylase activity (Govindaraj and Poulos, 1995, 1996).

Robin et al., fused the P450cam domain from *Pseudomonas putida* to the reductase domain of P450RhF from Rhodococcus sp. Strain NCIMB 9784 with linker lengths varying from 1 to 7 amino acids, named P450cam-RhFFRed L1-L7. No conversion of camphor (final concentration used = 3 mM) to 5-*exo*-hydroxycamphor was detected for P450cam-RhFFRed L1. P450cam-RhFFRed L2, L3 and L4 showed a conversion of 5 ± 2 % and 3.5 ± 1 % respectively after 24 h of reaction using whole cells. The conversion rate increased with the increase in length of the linker with highest product formation observed for P450cam-RhFFRed L7 showing a conversion of 83 ± 11.5% and 100% for reactions after 24 h and

48 h respectively (Robin et al., 2009). The study thus showed that the linker length is of crucial significance in determining the activity of a fusion protein.

Owing to the similarity in redox partners between P450 BM3 and P450cin, we anticipated that the orientation of CinA and CinC should be similar to that of P450 and reductase partners in P450 BM3. Thus the C terminal of CinA was fused to the N terminal of CinC and a length of upto 16 amino acids residues were chosen for insertion.

In order to predict the linker sequence to be inserted between the CinA and CinC domains, secondary structure prediction of the reported P450 BM3 linker (PSPSTDQSAKKVRKKA) was performed using the Chou-Fasman parameters (Chou and Fasman, 1974a, b).

Table 4.2.3. Secondary structure predictions of the linkers using the Chou-Fasman parameters (http://www.biogem.org/tool/chou-fasman/). The percentage possibility for secondary structure of each linker sequence has been classified in Helix, Sheet and Turn.

Linker	Sequence	Probability of sequence to form (%)		
		Helix	Sheet	Turn
BM3	PSPSTDQSAKKVRKKA	62.5	0.0	25.0
First eight amino acids of P450 BM3	PSPSTDQS	0.0	0.0	37.5
Last eight amino acids of P450 BM3	AKKVRKKA	87.5	0.0	25.0
Last eight amino acids new designed linker	PSTGDAVA	0.0	0.0	12.5
New designed linker	PSPSTDQSPSTGDAVA	0.0	0.0	25

The prediction showed a 62.5% probability of α-helix formation; the last eight amino acids of the P450 BM3 linker (AKKVRKKA) had an 87.5% probability of α-helix formation whereas the secondary structure formation probability of the first eight amino acids (PSPSTDQS) is only 37.5% (Table 4.2.2). To avoid introducing a linker prone to form secondary structures, the sequence of the first eight amino acids of the P450 BM3 were selected for CinA-CinC linker generation. The sequence of the subsequent eight amino acids was chosen (SPSTGDAVA) by analyzing frequently occurring sequences in natural fusion proteins which did not generate any secondary structure (George and Heringa, 2002; Wriggers et al., 2005). The chosen amino acid sequence for the full linker length of 16 amino acids between CinA and CinC was Pro-Ser-Pro-Ser-Thr-Asp-Glu-Ser-Pro-Ser-Thr-Gly-Asp-Ala-Val-Ala (PSPSTDQSPSTGDAVA).

P-Link can also be employed for insertion of linker lengths longer than 16 amino acid residues. It can be assumed that the efficiency of the insertion will not decrease since the insertion of linkers depend primarily on the cleavage of phosphorothioate bonds in amplified products using Iodine-ethanol and is not an enzymatic reaction.

Additionally P-Link can also be employed to vary the amino acid sequence within linkers (e.g. two positions with NNK codons) in the primer sequence can be incorporated to explore and tune flexibility of the linker sequence and thereby increase activity of the fusion protein. In this case, one suitable high throughput screening campaign based on flow cytometry (Ruff et al., 2012) or even based on agar based screening systems can be employed (Çelik et al., 2005) to screen the fusion protein linker library to identify the active constructs.

Slessor et al., reported construction of vectors for the polycistronic coexpression of P450cin and cindoxin (CinC) with or without terminal redox partner (*E. coli* Fpr). *E. coli* Fpr is used in P450cin mediated biocatalytic reactions since its natural redox partner CinB cannot be expressed in an active form. However, when Fpr was coexpressed with P450cin and cindoxin, the expression levels of cindoxin were drastically decreased presumably due to competition between cindoxin and Fpr for flavin cofactors. The coexpression system also decreased the expression levels of P450cin (Slessor et al., 2012). Whereas the fusion protein proposed in this thesis posed no difficulties in expression levels, neither for CinA or CinC. Use of fusion protein not only facilitates the expression but also simplifies P450cin system on genetic level making it suitable for directed evolution experiments.

Structural studies of linker regions of fusion proteins are very challenging. The length, flexibility, sequence of amino acids can have dramatic effects on the activity and structural stability of the fused domains. Given their high flexibility, loops tend to have an ambiguous conformation in crystal structures and their behavior and interactions with the fused domains are difficult to predict. Therefore, optimal linker lengths for protein fusion are difficult to design in a rational manner, and usually a single length is probed and used for generating fusion proteins, dramatically reducing their success rate. P-Link gives the possibility to generate a library of variable linker lengths increasing the likelihood to obtain an active fusion protein.

The direct fusion CinA-CinC (no linker) showed no measurable conversion of 1, 8-cineole. The latter is also the case for fusion proteins with a linker length of four and shorter however an equimolar mixture of unfused CinA and CinC show activity in presence of *E. coli* Fpr and NADPH (Table 4.2.1). One possible reason for the inactivity might be due to the misfolding of CinA or CinC or both CinA and CinC. Linker lengths less than 5 amino acids, possibly introduces restriction of movement of individual or both of the fused protein domains (results presented and discussed in chapter 4; section 4.3). This restriction on the movement of fused domains can limit efficient electron transfer in turn affecting the hydroxylation activities. However, at the current stage, without a crystal structure it is difficult to speculate the actual reason for the trend observed in hydroxylation activities of linker variants. A more detailed characterization in combination with computational and crystallographic studies might possibly give some insights regarding the observed activity trend.

4.2.9 Conclusion

In summary, a robust, technically simple and enzyme free method (P-LinK-Protein fusion with variable linker insertion) was developed for generating fusion proteins with variable linker length. The generation of fusion proteins and insertion of inter-domain linker of variable length can be achieved in a single experiment and screened for optimal properties, such as expression or enzymatic activity. CinC was fused to the C terminal end of CinA through sixteen different linker lengths in a single experiment employing three PCRs. The resulting CinA-CinC linker library was screened using a TLC based screening system to identify the most active CinA-linker10aa-CinC construct.

Although the fusion protein generated in this study shows lower activity compared to equimolar mixture of its individual component, the use of fusion proteins simplifies the multicomponent system on genetic level and enables its use in protein engineering campaigns. The latter can be used to obtain fusion protein variants with improved activity or with non-natural enzymatic properties.

P-LinK is a straight forward approach which minimizes experimental efforts and reduces significantly time demands and is therefore technically simple and robust. It is further likely that P-LinK can be expanded iteratively to at least three component systems in which a

third electron transfer partner of monooxygenase systems can be fused. The method can also be used for cytochrome P450s where native electron transfer partners are unknown. P -LinK is not only limited to P450 enzymes and can be of general interest to generate fusion enzymes to perform, for instance, cascade reactions.

4.3 Preliminary Insights into Role of Linker in Cytochrome P450cin Fusion Protein

4.3.1 Aim of the Work

The CinA-CinC fusion protein with a linker length of ten amino acids (CinA-$_{linker10aa}$-CinC) was identified as the most active variant for hydroxylation of 1,8-cineole (Chapter 4; section 4.2.5.1). Screening of CinA-linker $_{(1-16)}$ - CinC library also revealed linker length of four or less amino acids was determined to be inactive. This part of the result section, addresses two important aspects of the CinA-CinC fusion protein. (i) Role of length of linker sequence in modulating the activity of CinA-CinC fusion protein. (ii) The interaction of CinA and CinC in the fusion protein and its impact on the activity of the CinA-CinC fusion protein.

4.3.2 Construction of Homology Model of CinA - CinC Fusion Protein with Linker of Variable Length

Chapter 4, section 4.2 described the construction and activity determination of CinA and CinC fusion protein with a linker length of 1 to 16 amino acids. In order to visualize how the introduced linkers would translate into protein structure, homology models of the fusion and linker variants were generated using YASARA (Krieger et al., 2002). The template for construction of homology model was based on the coordinates of CinA (PDB id: 1T2B) (Meharenna et al., 2004). The crystal structure of CinC was not available during the time when this study was initiated. As a result a homology model of CinC was generated using YASARA. The template selected for construction of homology model of CinC was the FMN binding domain of human cytochrome P450 reductase (PDB id: 1B1C) (Zhao et al., 1999) with a sequence identity of 37%. Subsequently, CinA and CinC were used for modelling of linkers between the CinA-CinC fusion protein.

Madrona et al., recently reported the crystal structure of Cindoxin (CinC), the P450cin redox partner (PDB id: 4OXX) (Madrona et al., 2014). In order to evaluate if the homology model of the CinA-linker-CinC$_{(homology)}$§ constructed is similar to that of homology model of the

---

§ Homology model of CinA-linker-CinC constructed in YASARA using homology of CinC based on FMN binding domain of human cytochrome P450 reductase (PDB id: 1B1C) as template.

CinA-linker-CinC$_{(crystal)}$[**]; a homology model for CinA-$_{linker10aa}$-CinC$_{(crystal)}$ using the coordinates from crystal structure of CinC was constructed. The CinA-$_{linker10aa}$-CinC$_{(crystal)}$ homology model was compared to the CinA-$_{linker10aa}$-CinC$_{(homology)}$, however no significant difference were observed between the two models. Thus, the models generated using CinC$_{(homology)}$ were used for subsequent studies and analysis. The homology models of CinA - CinC fusion proteins with various linker lengths (6-16) are shown in figure 4.3.1.
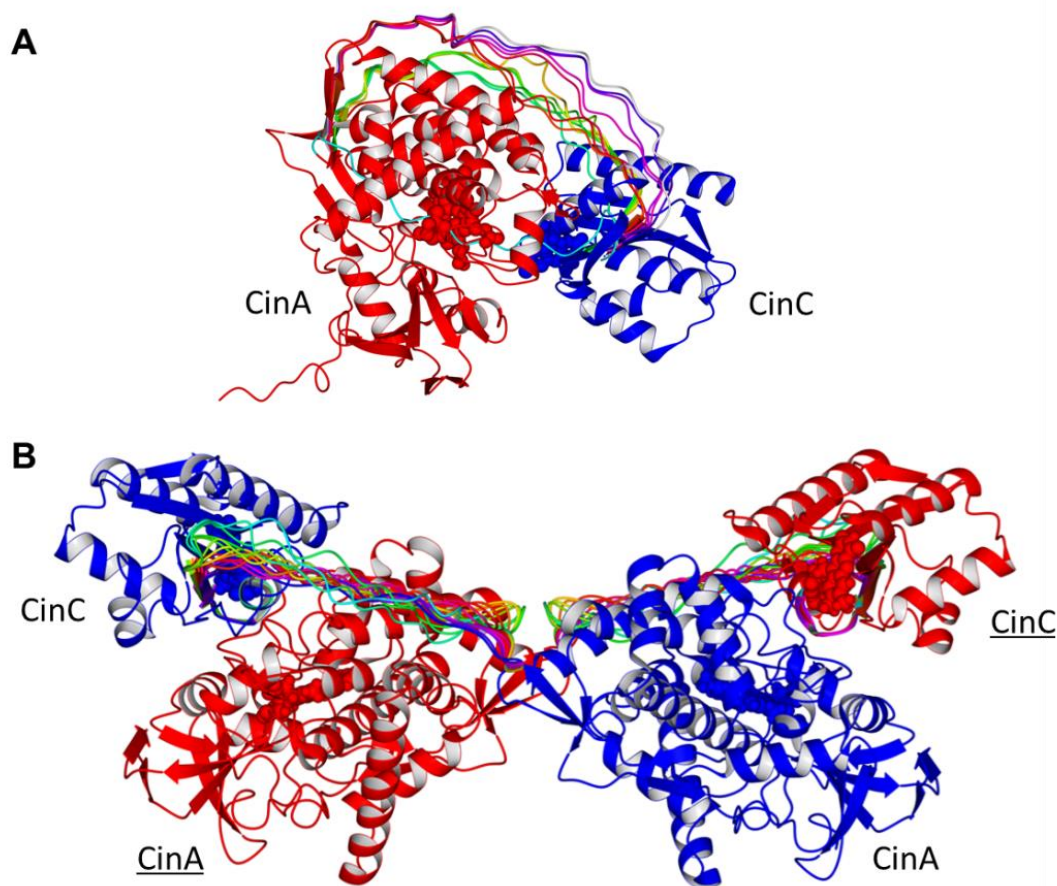

Figure 4.3.1: Hypothetical interactions and conformation between fused CinA and CinC. **A)** Intramolecular interaction. CinA forms a functional unit with the corresponding fused CinC. **B)** Intermolecular interaction. CinA forms a functional unit with a CinC domain from another fusion construct (blue), and conversely CinC from the first molecule (red) interacts with the remaining CinA. The modeled linkers of variable lengths for each configuration are represented in different color.

[**] Homology model of CinA-linker-CinC constructed in YASARA using coordinated from crystal structure of Cindoxin (CinC) (PDB id: 4OXX).

For linker lengths shorter than 6 amino acid residues it was difficult to model them in the fusion protein using YASARA. The length of the linkers was too short to give a correctly oriented and reliable homology model.

Interestingly, the resulting models indicated that the linkers identified as most active (9, 10, 11 amino acids long) between CinA and CinC would not allow the domains to achieve a proper orientation (based in P450 BM3 crystal structure) without significant structural changes (Figure 4.3.1). This suggests that the optimum linker length of ten amino acids might not be sufficient for domains from the same polypeptide chain to achieve electron transfer and an intermolecular electron transfer mechanism is plausible. Intermolecular electron transfer mechanism, meaning that a given heme domain would interact with a FMN domain coming from a different molecule and vice versa. Kitazume et al., proposed a similar mechanism for electron transfer in P450 BM3 (Kitazume et al., 2007) Experimental results based on site directed mutagenesis studies suggested that electron-transfer pathway in P450 BM3 passes through both protein molecules in the dimer during a single turnover, traversing from the FAD domain of one molecule into the FMN domain of the other molecule before passing to the heme domain (Kitazume et al., 2007).

4.3.3 Investigation of Intermolecular or Intramolecular Electron Transfer in CinA-CinC Fusion Protein

In order to elucidate if the P450cin fusion variants work as an intramolecular or as an intermolecular complementation system, we designed an experimental setup selectively knocking out the activity of individual domains as described below. Two SDM mutants of the CinA- linker10aa-CinC variant were constructed. The first variant (A91F) was designed to have an inactive heme domain (named as Heme (-)). The second variant (G594D) comprised an inactive FMN domain (named as FMN (-)).

The amino acid substitution G570D leads to inactivation of FMN domain in P450 BM3 (Kitazume et al., 2007). A structural alignment of P450cam, P450cin and P450 BM3 was performed, identifying position G594 as the structural homolog in P450cin for the construction of FMN (-). In the Heme (-) variant, alanine 91 was replaced by a bulkier phenylalanine residue which hindered the access of 1,8- cineole to the enzyme active site, leading to an inactive enzyme variant.

4.3.4 Site Directed Mutagenesis for Construction of Heme (-) and FMN (-) Variants

Two step site directed mutagenesis protocol was employed for construction of Heme (-) and FMN (-) variant as described in chapter 3; section 3.2.5 employing primers as mentioned in Appendix; Table 3. PCR products were run on 0.8% agarose gel (Figure 4.3.2) to confirm the amplification of the gene. An expected size of 3.8 kbp was observed on the agarose gel.
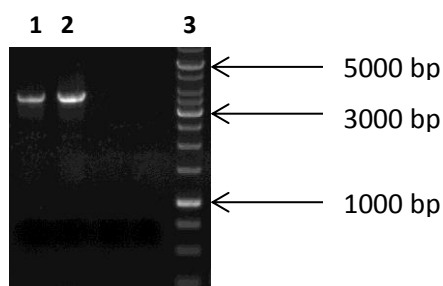

Figure 4.3.2. 0.8% agarose gel for site mutagenesis at position A91 and G594. Lane 1 = 1 kbp DNA ladder; lane 1 = SDM at position 81, lane 2 = SDM at position 594 (expected size ~ 3.8 kbp).

The amplified PCR products were digested and transformed in chemically competent *E. coli* BL-21 Gold (DE3) lacI^{Q1} cells as described in chapter 3. The mutation at sites A91 and G594 were confirmed by sequencing analysis. Expression and purification of the Heme (-) and FMN (-) variants was performed as described in chapter 3.

4.3.5. Determination of Activity of Heme (-) and FMN (-) Mutants to Analyze Electron Flow in CinA-CinC Fusion Protein

The activity of Heme (-) and FMN (-) variants were tested for activity towards 1,8-cineole. *E. coli* Fpr was used to reconstitute the system and NADPH (0.4 mM) was used as reduction equivalent using 1 μM P450 with a reaction time of 2 h. The reactions were performed as described in chapter 3, section 3.3.1.4 and the amount of 2-β-hydroxy-1,8-cineole formed was quantified using GC-FID as described in chapter 3, section 3.4.3.

Activity determination was performed for the flowing set of variants (i) Heme (-)-linker 10 aa-CinC; (ii) CinA-linker 10 aa-FMN (-) (iii) combination of Heme (-)-linker 10 aa-CinC

and CinA-linker 10 aa-FMN. All reactions were performed in triplicate. The activity profile for all the reactions is presented in table 4.3.1. The Heme (-)-linker 10 aa-CinC and CinA-linker 10 aa-FMN (-) were individually inactive towards 1,8-cineole (Table 4.3.1). However, the combination of purified Heme (-) and FMN (-) variants, however, showed a hydroxylation activity comparable to the CinA- linker 10 aa- CinC fusion protein, strongly supporting the presence of intermolecular interaction between CinA (Heme) and CinC (FMN). These experimental results support the theory based on computational analysis (Figure 4.3.1).

Table 4.3.1 2-β-hydroxy-1,8-cineole formation for Heme (-) and FMN (-) mutants. Heme (-) mutant had mutation A91F and FMN (-) mutant had mutation G594D individually in heme and CinC domain in CinA-linker10aa-CinC fusion protein. Reactions were performed with *E.coli* Fpr, NADPH [final concentration = 0.4 mM] and 1 µM P450 for a reaction time of 2 h. All reactions were performed in triplicate. n.d. = no product detected.

Variant	Amount of 2-β-hydroxy-1,8-cineole formed[µM µM P450^{-1}]
CinA-linker 10aa-CinC	176.7±2.4
Heme(-)-linker 10 aa-CinC	n.d.
CinA-linker 10 aa-FMN(-)	n.d.
Heme(-)-linker 10 aa-CinC + CinA-linker 10 aa-FMN(-)	174.0±8.7

An illustration of plausible electron transfer pathway in CinA-CinC fusion protein is shown in Figure 4.3.3. Shown in this figure, the FMN (-) domain was not able to transfer electrons from NADPH to CinA (heme domain) for conversion of 1,8-cineole to 2-β-hydroxy-1,8-cineole (Figure 4.3.3 B). Similar is the case with heme (-) domain where no product formation was detected (Figure 4.3.3 A). However, when the Heme (-)-linker 10 aa-CinC was combined with CinA-linker 10 aa-FMN (-), it showed a hydroxylation activity comparable to the CinA- linker 10 aa- CinC fusion protein. Figure 4.3.3 C shows the plausible electron transfer pathway from FMN of one molecule to the heme of another molecule strongly supporting inter-molecular electron transfer in P450cin fusion proteins. These results substantiate the theory based on computational analysis for CinA-linkers (6-16)-CinC constructs.
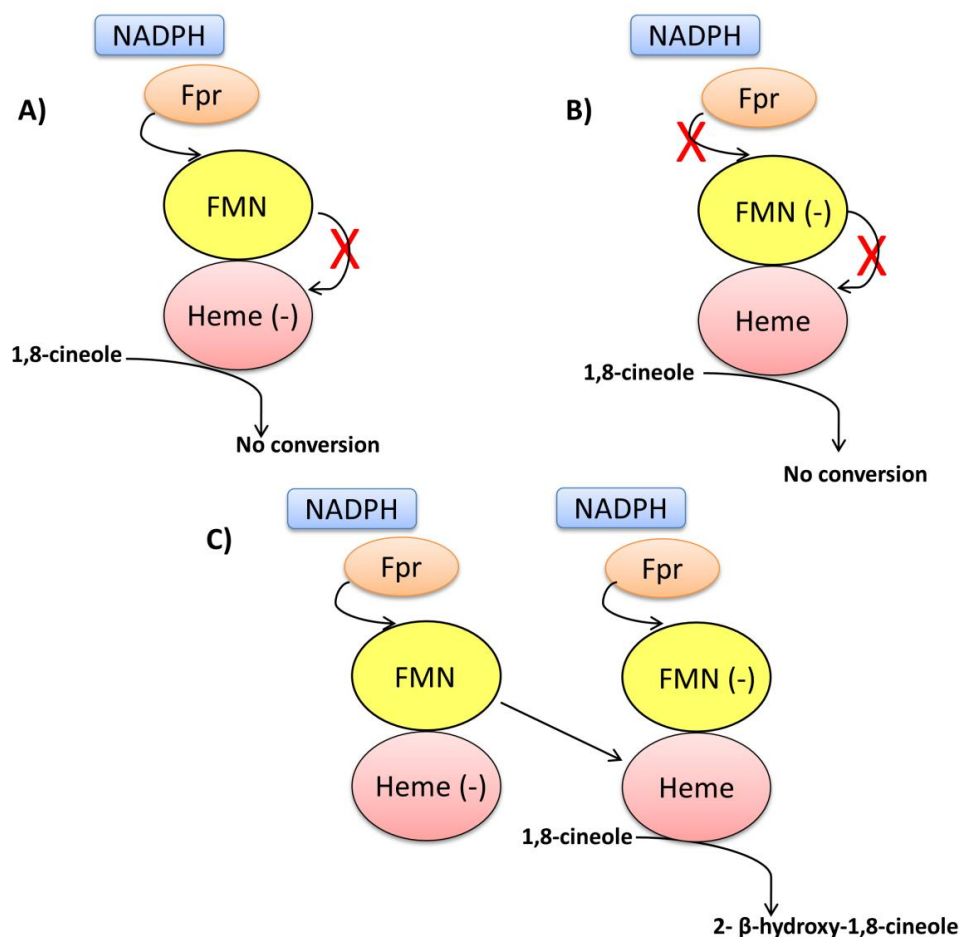
Figure 4.3.3 Illustration of electron transfer pathway in CinA-CinC fusion protein. Domains are arranged in order of reduction potential with most positive on top (FAD) followed by FMN and most negative at the bottom (P450). Alignment of the two domains is not intended to imply structural relationships but is consistent with the data obtained. **(A)** Heme (-) carries mutation A91F in the heme domain (CinA domain). **(B)** FMN (-) carries mutation G594D in the FMN domain (CinC domain). **(C)** Heme (-) mutant and FMN (-) mutant were mixed together in a single reaction tube to determine the activity. It also shows plausible intermolecular electron transfer pathway. Fpr = *E. coli* Flavodoxin reductase, FMN = CinC domain, Heme = P450cin. Arrows show possible electron transfer pathway. Cross represents no electron transfer.

The data obtained also suggests the possibility of presence of dimeric structure for P450cin fusion protein with the transfer of electrons from one monomer to another monomer. However, additional experiments to support the presence of dimeric structure for P450cin fusion protein were not performed and hence it can only be speculated with the data available. Studies based on determination of molecular mass, cross linking studies using

glutaralehyde or a simple native PAGE experiment might provide indication regarding the dimeric or monomeric form of the P450cin fusion protein.

4.3.6. Investigation of Role of Linker Length on the Activity of CinA-CinC Fusion Protein

Using the evidence supporting intermolecular activation provided by the experimental data, a remaining question regarding the inactive linker variants was addressed. For CinA-CinC fusion protein linker lengths shorter than 4 amino acids were inactive with no detectable conversion of 1,8-cineole to 2-β-hydroxy-1,8-cineole.

In order to determine if the fusion proteins with linker length 4 or less are inactive due to misfolding or due to the inability of CinA to interact in the right configuration with the fused CinC, the inactive linker variants were supplemented with 1 μM CinA or CinC, *E. coli* Fpr and NADPH [final concentration = 0.4 mM] and product formation and coupling were measured using GC- FID measurements as described previously in chapter 3, section 3.4.3. Table 4.3.2 shows the amount of 2-β-hydroxy-1,8-cineole formed by supplementing the inactive linker variants (direct CinA-CinC fusion and CinA- CinC with linker lengths of 1 – 4) with 1 μM of CinA or 1 μM of CinC.

Interestingly, all inactive fusion variants (direct CinA-CinC fusion and CinA - CinC with linker lengths of 1 – 4) recovered activity by either supplementing CinA or CinC (Table 4.3.2), suggesting that in all tested constructs, the CinA and CinC partners are folded in an active form. The direct fusion, however, only recovered approximately 33 % of the activity of the reconstituted unfused CinA + CinC system, whereas the CinA-CinC fusion proteins with linker 1 to 4 recovered approximately 95% of the reconstituted unfused system (Table 4.3.2).

By observing these results, it is clear that there are two main parameters governing the activity of the different fused partners. The first is the correct folding of each individual partner when fused, which seems to be, by the results of the above activity assay, correct in all investigated variants, even in the direct fusion of CinA and CinC. The second parameter and the more crucial one, is the length of the linker, which appears to govern the activity of the fused partners. The difference in the activity of the CinA-CinC fusion proteins can be concluded to be due to the difference in the linker length.

Table 4.3.2. Activity profile for inactive linker variants (direct CinA-CinC fusion and CinA - CinC with linker lengths of 1 – 4). Reactions were performed with *E.coli* Fpr, NADPH [final concentration = 0.4 mM] and 1 μM P450, 1 μM CinC for a reaction time of 2 h. All reactions were performed in triplicate. n.d. = no product detected. Product formed refers to the formation of 2-β-hydroxy-1,8-cineole formed (μM μM^{-1} P450). The coupling efficiency (%) was calculated as the ratio of NADPH used for substrate oxidation to the total amount of NADPH consumed by P450 in the reaction.

Construct	Product formed	Coupling efficiency (%)	+ 1 μM CinA		+ 1μM CinC	
			Product formed	Coupling efficiency (%)	Product formed	Coupling efficiency (%)
CinA	n.d.	n.d.	n.d.	n.d.	300±13	75.0±3.4
CinC	n.d.	n.d.	299±11	74.8±3.5	n.d.	n.d.
CinA-0-CinC	n.d.	n.d.	106±11	26.6±2.8	105±11	26.5±2.8
CinA-1-CinC	n.d.	n.d.	292±2	73.1±0.3	293±0	73.3±0.0
CinA-2-CinC	n.d.	n.d.	268±4	67.0±1.1	270±6	67.6±1.4
CinA-3-CinC	n.d.	n.d.	269±7	67.3±1.7	272±2	68.2±0.4
CinA-4-CinC	n.d.	n.d.	250±11	62.5±2.9	257±11	64.2±2.6
CinA-5-CinC	22.0±3.2	5.5±0.1	246±6	61.5±1.2	247±5	61.7±1.3

4.3.7 Discussion

Homology modelling of CinA-linker-CinC proteins (described in section 4.3.2) revealed that a linker length of 9 - 11 amino acids which was determined to be optimum for hydroxylation of 1,8-cineole might not sufficient for electron transfer between heme and FMN domains form the same polypeptide chain. In this case there are two possibilities, one the electron transfer will be between two molecules (intermolecular) or P450cin exists as a dimer and electron transfer is between the dimers, in a similar way as reported for P450 BM 3.
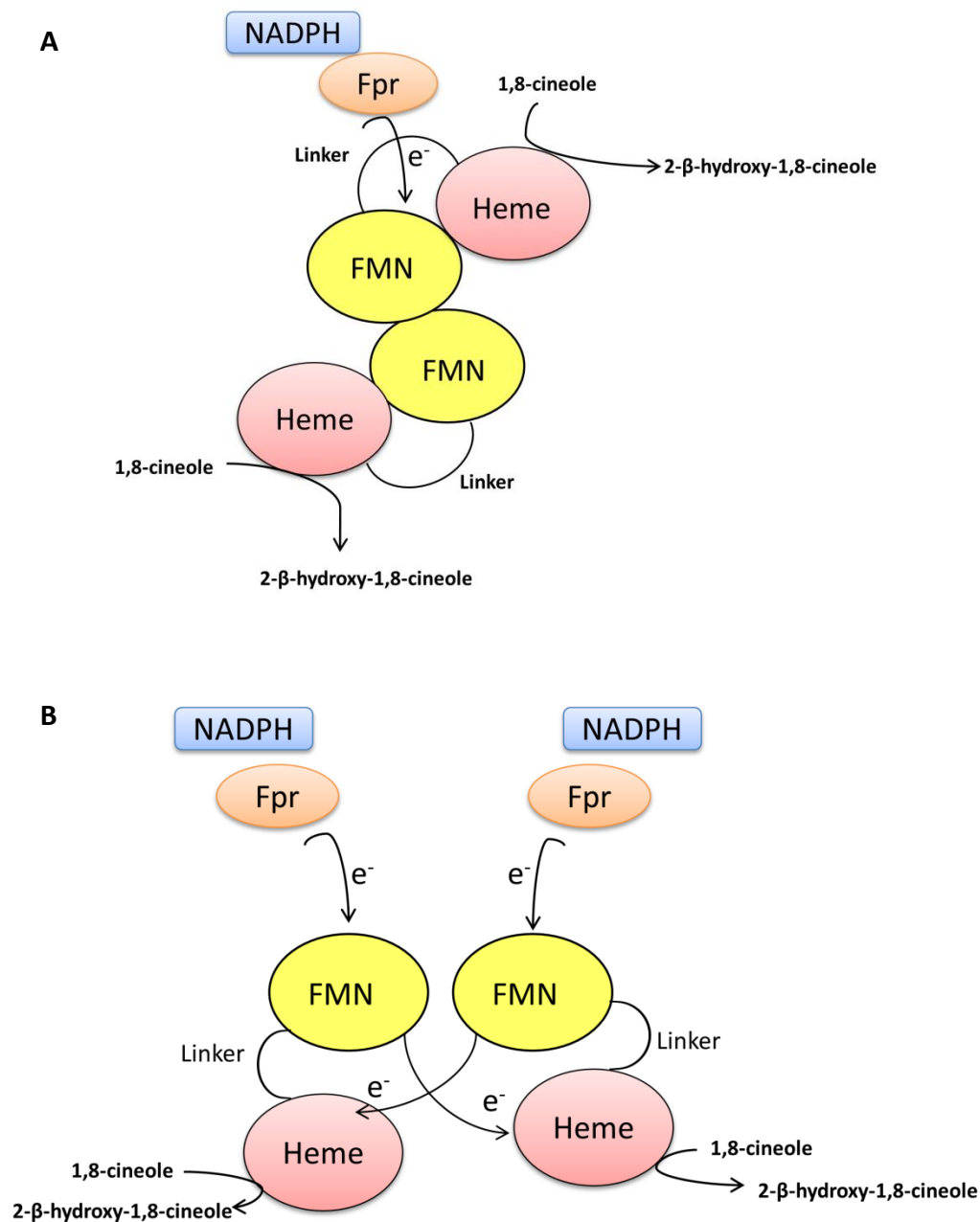
Figure 4.3.4. (A) and **(B)** represent the hypothetical model for electron transfer pathway in P450cin fusion protein. **(A)** Putative dimeric form of P450cin fusion protein. The electron transfer pathway in dimeric form might occur between the FMN of one monomer to the heme of another monomer. **(B)** Putative model for intermolecular electron transfer pathway in P450cin fusion protein.

Kitazume et al proposed a mechanism for electron transfer in P450 BM3. P450 BM 3 exists as a dimer and experimental results based on site directed mutagenesis studies suggest that electron-transfer pathway in P450 BM3 passes through both protein molecules in the dimer during a single turnover, traversing from the FAD domain of one molecule into the FMN domain of the other molecule before passing to the heme domain (Kitazume et al., 2007). In order to determine if the electron transfer pathway in P450cin is similar to that observed in P450 BM 3, two site directed mutants were generated to knock out the activity of individual domains. The individual inactive domains (Heme (-) and FMN (-)) were found to be inactive and no hydroxylation activity was detected. However, for an equimolar mixture of individual inactive domains (Heme (-) + FMN (-)) formation of 2-β-hydroxy-1,8-cineole was detected. Figure 4.3.4 shows the hypothetical models for electron transfer pathways in P450cin fusion protein supporting intermolecular electron transfer. In case of mammalian fusion enzyme nitric oxide synthase (NOS), the catalytically active site form of the enzyme is a homodimer with dimerization solely via the heme binding oxygenase domain, where as in case of P450 BM3, the protein dimerizes via the reductase domain. With the experimental evidence obtained for the study in this thesis for P450cin it is however difficult to speculate if P450cin exists as a dimer or a monomer. Additional characterization of the fusion protein with electrospray ionization mass spectroscopy (ESI-MS), using nuclear magnetic resonance (NMR), cross linking studies with glutaralehyde, fluorescence anisotropy, etc. will provide further evidence if the P450cin fusion protein exists as a monomer or a dimer.

Another interesting observation from the experimental result was that the linker length was of crucial significance in the activity of the fusion protein. The role of the linker region is not clear however, it is commonly postulated that the linker region acts as a flexible hinge allowing recognition and interaction between the two domains (Hannemann et al., 2007; Sevrioukova et al., 1999). Previous reports demonstrate that the length of the linker is decisive for activity of fusion proteins rather than the amino acid composition of the linker. In two individual studies Govindraj and Poulos reported that changing Arg-Lys-Lys stretch in the middle of the linker sequence in P450 BM3 to Ala-Ala-Ala does not alter the functional properties of either of the two domains. However, deleting three or six amino acids leads to drastic effect of the hydroxylation activity of enzyme. The three amino acid deletion mutant exhibited 10 % of the hydroxylation activity while the six amino acid

deletion mutant showed nearly undetectable levels of fatty acid hydroxylase activity (Govindaraj and Poulos, 1995, 1996). In a similar study by Robin et al., where the P450cam domain from *Pseudomonas putida* was fused to the reductase domain of P450RhF from Rhodococcus sp. Strain NCIMB 9784 with linker lengths varying from 1 to 7 amino acids, named P450cam-RhFFRed L1-L7.(Robin et al., 2009). No conversion of camphor (final concentration used = 3 mM) to 5-*exo*-hydroxycamphor was detected for P450cam-RhFFRed L1. The conversion rate increased with the increase in length of the linker with highest product formation for P450cam-RhFFRed L7 (Robin et al., 2009).

Similar results were obtained with the CinA-CinC fusion protein where increase in hydroxylation activity was observed with the increase in linker length. The lower length linker (length less than 4 amino acids) however were inactive. External addition of cell lysate of CinA and CinC to inactive linkers showed hydroxylation of 1,8-cineole to 2-β-hydroxy-1,8-cineole. Thus, the folding of CinA and CinC were in the correct orientation however, the length of the linkers was not sufficient to hold the fused CinA and CinC domains in the required orientation.

4.3.8 Conclusion

To conclude, experimental evidence from our study suggests that the electron transfer in CinA-CinC fusion protein predominantly is intermolecular. The plausible electron transfer pathway is from FMN domain of one molecule to the heme domain of another molecule. Additional investigations show that the CinA and CinC domain in the fusion protein are correctly folded and the length of the linker separating the fused domains is crucial for activity of the fusion protein.

4.4 Directed Evolution of P450cin Fusion Protein for Mediated Electron Transfer

4.4.1 Aim of the Work

Reactions catalyzed by P450 monooxygenase are *stereo-* and *regio-* selective and can be performed in water, at room temperature and under atmospheric pressure which is often difficult to achieve using chemical catalysts. One of the main hurdle for catalytic applications of cell free P450 catalyzed reactions is the requirement of NAD(P)H, as a reduction equivalent for each catalytic cycle. Using random protein engineering approach cytochrome P450cin fusion protein (CinA-linker10aa-CinC) was evolved for mediated electron transfer using Zinc and Cobalt(III)sepulchrate[††] as an alternative and cost-effective cofactor system.

4.4.2 Generation of Random Mutagenesis Libraries

Two iterative rounds of random evolution employing SeSaM were performed to identify the residues which modulate the activity of cytochrome P450cin towards 1,8-cineole while using alternative cofactor system Zn/Co(III)sepulchrate. The first round of random mutagenesis was performed with low mutational load as described in chapter 3 section 3.2.4. The template used for construction of SeSaM library was a fusion protein variant identified as the most active fusion variant identified for hydroxylation of 1,8-cineole (CinA-linler10aa-CinC) (Chapter 4, section 4.1.5.1). Template for SeSam library II was a double mutant Q385H/V386S, an improved variant identified from the first SeSaM library.

Electrophoretic analyses of steps involved in SeSaM technique used for mutant library generation are shown in figure 4.4.2. Preliminary step was performed for generation of template for step 1 and step 3 (Figure 4.4.2 A). The concentration of dATPαS [A forward- 30%, G forward- 25%] and dGTPαS [A reverse- 30%, G reverse- 25%] were selected in step 1 of SeSaM based on the fragmentation pattern (Figure 4.4.2 B and C). The selected concentrations were used in step 1 for construction of both SeSaM libraries. Step 2 is the TdT catalyzed incorporation of dPTαP at the 3'-OH of a 5'-FITC labelled oligonucleotide. The

number of incorporated P nucleotides was analyzed on a 36% acrylamide gel and visualized using a phosphorimager. Figure 4.4.2 D shows the TdT catalyzed incorporation of dPTPαS. Figure 4.4.2 E shows the amplification of gene after step 4 of SeSam. The PCR products obtained after step 4 were cloned in pALXtreme-1a by PLICing (50 ng each)(described in chapter 3; section 3.2.2) and transformed in chemically competent *E. coli* BL-21 Gold (DE3) LacI[Q1] cells (chapter 3; section 3.1.4).
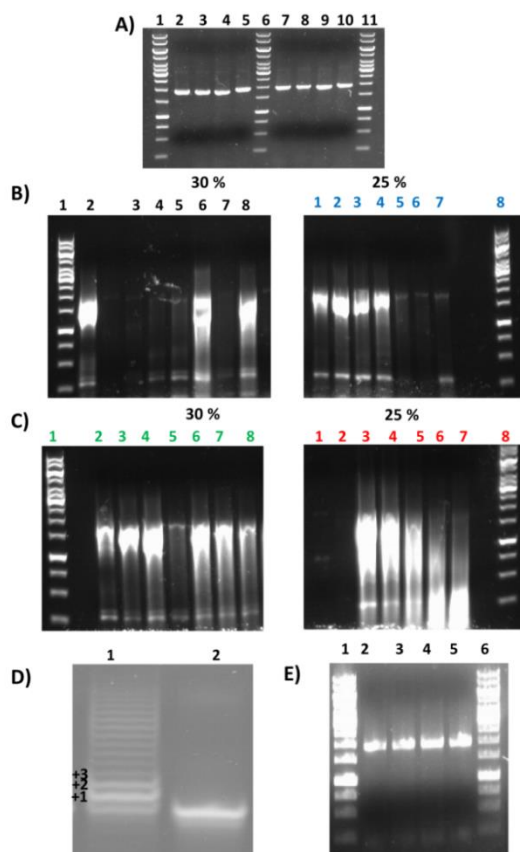

Figure 4.4.1. Electrophoretic analyses of steps in SeSaM technique used for mutant library generation. **(A)** Preliminary step- PCR amplification to generate template for step 1 and step 3. Lane 1= 1 kbp DNA ladder. **(B)** and **(C)** Determination of optimal phosphorothioate percentage using different concentrations of dATPαS or dGTPαS (lane 2 = 10, lane 3 = 15, lane 4 = 20, lane 5 = 25, lane 6 = 30, lane 7 = 35, lane 8 = 40 percent) for A_forward library (black), G_forward library (blue). (lane 1 = 10, lane 2 = 15, lane 3 = 20, lane 4 = 25, lane 5 = 30, lane 6 = 35, lane 7 = 40 percent, lane 8 = 1 kbp DNA ladder) for A_reverse library (green) and G_reverse library (green). **(D)** Phosphorimager analysis of 3'-tailed FITC oligonucleotide separated on 36 % acrylamide gel. Lane 1 = FITC catalyzed stepwise incorporation (+1, +2, +3) of dPTαP; lane 2 = FITC only (negative control). **(E)** Step 4 of SeSam; replacement of universal bases; L = 1 kbp DNA ladder; lane 1= A_forward library; lane 2 = G_forward library; lane 3: A_reverse library; lane 4 = G_reverse library.

Table 4.4.1: Sequencing analysis of random clones from SeSaM library I and SeSaM library II

A) SeSaM Library I

Mutant	C1	C2	C3	C4	C5
Nucleotide changes	CT->AC	GT->TC T->C T->G	C->T	C->T G->T	-
Amino acid changes	L520T	V417S S439S S445S	A53V	L357F	-
Position of mutation	CinC domain	CinA domain	CinA domain	CinA domain	
No. of nucleotide changes	2	4	1	2	-
Single Ts	0	1	1	1	-
Single Tv	0	1	0	1	-
TvTs	1	1	0	0	-
Deletions	0	0	0	0	-

B) SeSaM Library II

Mutant	C1	C2	C3	C4	C5
Nucleotide changes	G->T	G->T	G->T	-	-
Amino acid changes	G57V	E113STOP	W46C	-	-
Position of mutation	CinA domain	CinA domain	CinA domain	-	
No. of nucleotide changes	1	1	1	-	-
Single Ts	1	1	1	-	-
Single Tv	0	0	0	-	-
TvTs	0	0	0	-	-
Deletions	0	0	0	-	-

Quality of the generated mutant library was verified by sequencing of five random clones. Sequencing revealed 60% transitions (Ts), 20% transverion followed by transitions (TvTs;

subsequent mutation) and 20% wild type for SeSaM library I (Table 4.4.1 A) and 60% transitions (Ts) and 40% wild type for SeSaM library II (Table 4.4.1 B). Diversity generated in the both the SeSaM libraries as reveled by sequencing analysis matched the requirement of a mutant library and hence were used for subsequent screening.

4.3.3 Screening of SeSaM library I

Screening of approximately 2700 clones using TLC based screening system (described in chapter 4; section 4.1.4.3) for detection of 2-β-hydroxy-1,8-cineole was performed using NADPH as reduction equivalent. The activity assay was performed as described in chapter 3; section 3.3.1 using crude cell lysates (prepared as described in chapter 3; section 3.2.9). Subsequently 200 clones were selected which showed activity higher than 20% compared to wild type for rescreening. Following rescreening 16 clones were selected and expressed in expression flask (200 ml) and reactions for these clones were performed using crude cell lysate using normalized amount of P450 in the reaction. 2-β-hydroxy-1,8-cineole formation was quantified using GC-FID for the selected 16 clones of which four clones showed improvement in activity compared to wild type (CinA-linker10aa-CinC). Two out of the four selected were wild type enzyme and the other two variants carried single mutation each. Screening of first round of SeSaM library yielded two improved mutants with single substitution each (Q385H and V386S). KB1 (Q385H) and KB2 (V386S) showed a 1.5 and 2 fold improvement in hydroxylation activity as compared to the wild type variant (CinA-linker10aa-CinC) while using NADPH as a reduction equivalent. (Figure 4.4.2, Table 4.4.2). Variants KB1 and KB2 identified from SeSaM library round I were subsequently tested for activity using Zinc dust and Cobalt (III) Sepulchrate (described in chapter 3, section 3.3.1.4). The amount of 2- β-hydroxy-1,8-cineole formed for variants KB1, KB2 and WT (CinA-linker 10aa- CinC) is shown in Table 4.4.2 and Figure 4.4.3. The improvement in activity observed with Zinc dust and Cobalt(III)sepulchrate was however only 1.5 and 1.2 fold for variants KB1 and KB2 respectively (Figure 4.4.3 and Table 4.4.2).
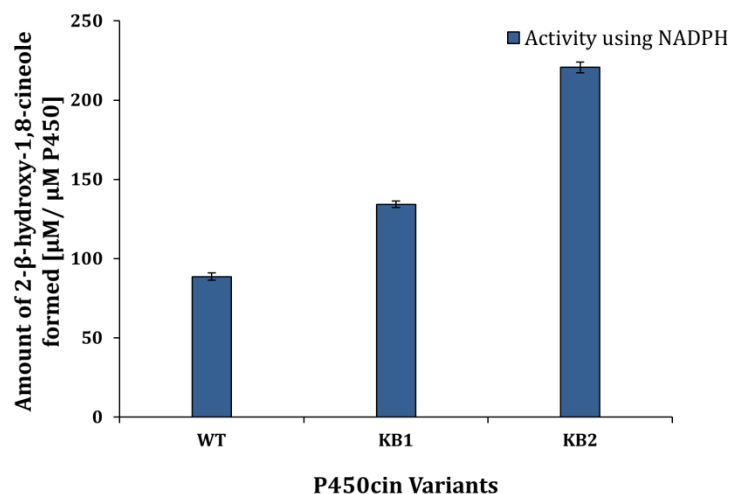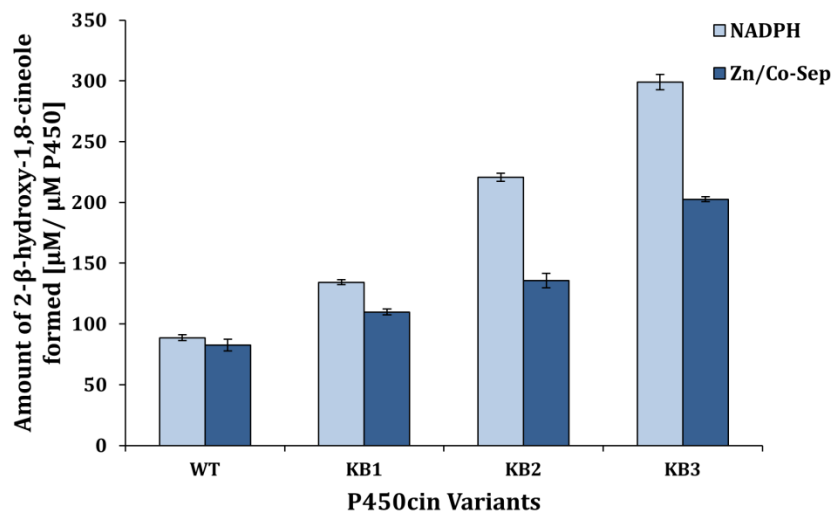
Figure 4.4.2. 2-β-hydroxy-1,8-cineole formed for P450cin variants identified from SeSaM library round I. KB1 variant (Q385H) and KB2 variant (V386S) showed a 1.5 and 2 fold improvements in activity with NADPH (0.4 mM final concentration) as reduction equivalent.

Table 4.4.2. Activity profile for variants KB1 (Q385H) and KB2 (V386S) identified after SeSaM round I. Reactions using NADPH [0.4 mM] as well as reactions using Zinc dust [5 mg] and Cobalt(III)sepulchrate [5 mM] were performed for 2 h at 30 °C using 1 μM P450. All reactions were performed in triplicate. Reaction products were extracted with ethyl acetate and analyzed by GC-FID for determination of 2-β-hydroxy-1,8-cineole.

Variant	2-β-hydroxy-1,8-cineole with NADPH ($\mu M\ \mu M^{-1}$ P50)	2-β-hydroxy-1,8-cineole with Zinc dust and Cobalt(III)sepulchrate ($\mu M\ \mu M^{-1}$ P50)
WT	88.67± 2.36	82.67 ± 4.94
KB1	134.30 ± 1.95	109.87 ± 2.38
KB2	220.67 ± 3.48	135.70 ± 6.06
KB3	298.89 ± 6.30	202.59 ± 1.95

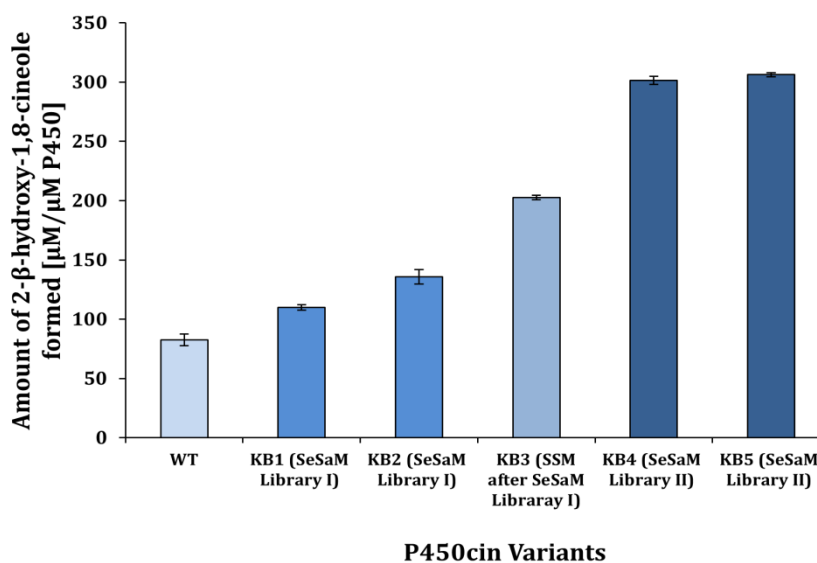**Figure 4.4.3** 2-β-hydroxy-1,8-cineole formed for P450cin variants identified from SeSaM library round I (KB1-Q385H and KB2- V386S) and KB3 (Q385H/V386S) identified from site saturation library at position Q385 and V386. All measurements were performed in triplicate using NADPH [0.4 mM] or using Zinc dust [5 mg] and Cobalt(III)sepulchrate [5 mM] and were performed for 2 h at 30 °C using 1 μM P450.

Though the improvement in activity of P450cin mutants identified in SeSaM library round I was not as expected for alternative cofactor system Zinc and Cobalt(III)sepulchrate nonetheless, the amino acid positions identified were used for generation of a site saturation mutagenesis library (see section 4.4.3.1). The primary reason for difference in improvement of activity for variants KB1 and KB2 using NADPH as electron source and alternative electron transfer system Zn/Cobalt(III)sepulchrate is due to the fact that the screening was performed using NADPH and not Zn/Cobalt(III)sepulchrate. It is an established fact in directed evolution campaigns "you get what you screen for" (You and Arnold, 1996).

4.4.3.1 Site Saturation Mutagenesis at Positions Q385 and V386

The two positions identified (Q385 and V386) in the first round of random mutagenesis were subjected to simultaneous site saturation mutagenesis in order to identify the most beneficial substitution at each position using NNK codon degeneracy in the primer sequence (Appendix, Table 4). However, screening of approximately 450 clones using TLC

based screening system and NADPH as reducing equivalent (see Chapter 3, Section 3.3.1.4) could not yield a mutant which was more active than combination of individual mutations (Q385H/V386S) (Table 4.4.2, Figure 4.4.3). This double mutant (Q385H/V386S or KB3) was subsequently used as a template for generation of round II of SeSaM.

4.4.4 Screening of SeSaM library II

Screening was performed for approximately 1350 clones using Zinc dust and Cobalt(III)Sepulchrate as electron transfer system and TLC screening system using crude cell lysate (as described in Chapter 3, Section 3.3.1.4). 100 clones showing an improvement in 2-β-hydroxy-1,8-cineole formation by 20% compared to WT enzyme (Q385H/V386S) were selected for re-screening. Re-screening was performed in a 96-well deep well plate using crude cell lysate and amount of 2-β-hydroxy-1,8-cineole formed was semi-quantified using TLC screening system. Subsequently, reactions with normalized amount of P450 was performed for 15 best clones and the amount of 2-β-hydroxy-1,8-cineole formed was quantified using GC-FID.

Figure 4.4.4. Activity profile for P450cin variants identified from SeSaM library round II (KB1-Q385H and KB2- V386S) and mutant KB3 (Q385H/V386S) identified from Site saturation library at position Q385 and V386. Variants KB4 (Q385H/V386S/T77C) and KB5 (Q385H/V386S/L88N) were identified from SeSaM library round II. All measurements were performed in triplicate using Zinc dust [5 mg] and Cobalt (III) Sepulchrate [5 mM] and were performed for 2 h at 30 °C using 1 μM P450.

Screening yielded two improved variants with single mutation each (KB4 and KB5). Variants KB4 (Q385H/V386S/**T77C**) and KB5 (Q385H/V386S**L88N**) exhibited a 3.5 fold improvement in hydroxylation activity in comparison to WT (CinA-linker 10aa-CinC) (Figure 4.4.4).

4.4.4.1 Site Saturation Mutagenesis at Positions T77 and L88

The amino acid positions identified from SeSaM library round II were subjected to simultaneous site saturation mutagenesis. PCR was performed (described in chapter 3; section 3.2.5) using mixture of four primers containing NNK degeneracy in the sequence of the forward primer (Appendix, Table 4). Five random clones were sequenced in order to determine randomization of amino acid residues at positions 77 and 88.
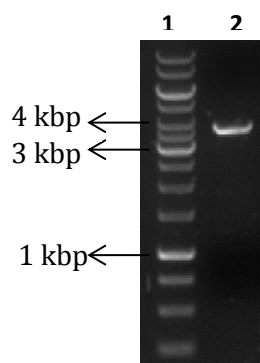
Figure 4.4.5. 0.8% agarose gel for simultaneous site saturation mutagenesis at position 77 and 88. Lane 1 = 1 kbp DNA ladder; lane 2 = PCR product for SSM at position 77 and 88 (expected size ~ 3.8 kbp).

Table 4.4.3. Sequencing analysis of random clones from simultaneous SSM library at position 77 and 88.

Clone	Position 77	Position 88
WT	acc	ctg
1	acc	att
2	cgt	act
3	act	acg
4	ccg	tcg
5	act	tgt

Sequencing analysis of five random clones showed randomization of amino acids at position 77 and 88 (Table 4.4.3). The simultaneous site saturation library was screened using Zinc/Cobalt(III)sepulchrate (described in chapter 3; section 3.3.1.4). Screening of about 450 clones was performed using the TLC based screening system. Mutants showing activity improvement by about 2 folds were selected for rescreening. The rescreening of the selected clones (30 clones) was performed using normalized amount of P450 (1 µM) in the reaction mixture and the formation of 2-β-hydroxy-1,8-cineole was determined using GC-FID measurements. Simultaneous SSM library generated three mutants KB6 (Q385H/V386S/**T77S/L88C**), KB7 (Q385H/V386S/**T77C/L88Y**) and KB8 (Q385H/V386S/**T77N/L88R**) with improvement in activity by about 5 folds (Figure 4.4.6) compared to WT (CinA-linker10aa-CinC).
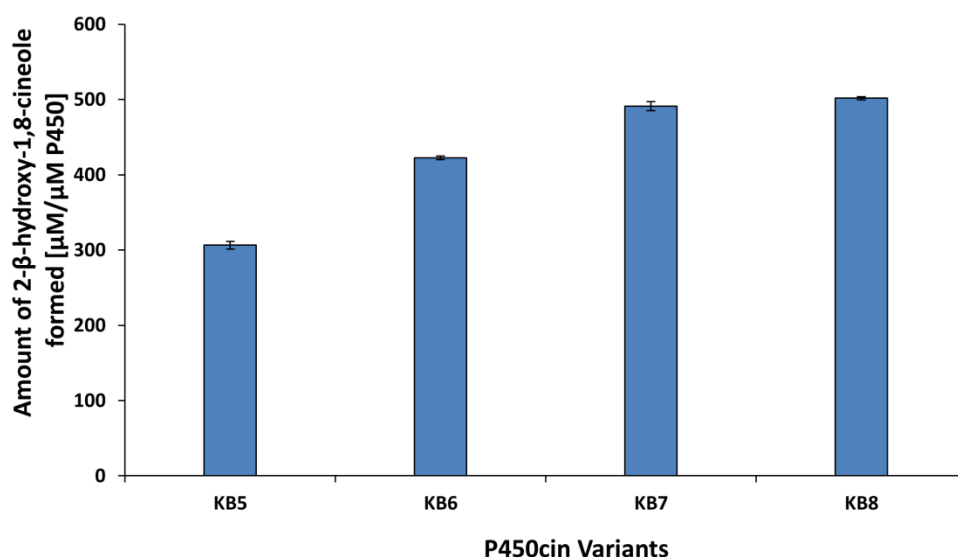
Figure 4.4.6. 2-β-hydroxy-1,8-cineole formed for P450cin variants identified after screening of SSM library at amino acid position 77 and 88. KB5 (Q385H/V386S/L88N) were identified from SeSaM library round II. Variant KB6 (Q385H/V386S/T77S/L88C), KB7 (Q385H/V386S/T77C/L88Y) and KB8 (Q385H/V386S/T77N/L88R) were identified from SSM library at amino acid position T77 and L88. All measurements were performed in triplicate using Zinc dust [5 mg] and Cobalt(III)sepulchrate [5 mM] and were performed for 2 h at 30 °C using 1 µM P450.

4.4.5. Comparing Activities of P450cin Variants with NADPH and Zn/Cobalt(III) sepulchrate

In order to compare the activities of P450cin variants identified after two iterative rounds of SeSaM for electron source Zn/Cobalt(III)sepulchrate and NADPH, amount of 2-β-hydroxy,1,8-cineole formed was determined for WT (CinA-linker10aa-CinC), variants KB1-KB8 and an equimolar mixture of individual component of P450cin system (CinA+CinC+Fpr).

Table 4.4.4 gives a comparison of the product formation of equimolar mixture of individual components of P450cin (CinA+CinC+Fpr) and KB8. Variant KB8 showed around 2.2 times increased 2-β-hydroxy-1,8-cineole formation (501.56 ± 9.19 μM μM P450^{-1}) using electron donor system Zn/Cobalt(III)sepulchrate compared to an equimolar mixture of CinA+CinC+Fpr (222.22 ± 3.23 μM μM P450^{-1}). In addition mutant KB8 displayed a 1.5 times improvement in activity (501.56 ± 9.19 μM μM P450^{-1}) using Zn/Cobalt(III)sepulchrate compared to the equimolar mixture of CinA+CinC+Fpr (315 ± 13.88 μM μM P450^{-1}) using NADPH as reduction equivalent (Table 4.4.4 and Figure 4.4.7).

Table 4.4.4. Activity profile for variant KB8 and equimolar mixture of individual components of P450cin (CinA+CinC+Fpr). Reactions using NADPH [0.4 mM] as well as reactions using Zinc dust [5 mg] and Cobalt(III)sepulchrate [5 mM] were performed for 2 h at 30 °C using 1 μM P450 in triplicate. Reaction products were extracted with ethyl acetate and analyzed by GC-FID for determination of 2-β-hydroxy,1,8-cineole.

Variant	2-β-hydroxy-1,8-cineole formed [NADPH] (μM μM P50^{-1})	2-β-hydroxy-1,8-cineole formed [Zinc dust and Cobalt (III) Sepulcharte] (μM μM P50^{-1})
CinA+CinC+Fpr	315 ± 13.88	222.22 ± 3.23
WT	88.67± 2.36	82.67 ± 4.94
KB8	680 ± 5.5	501.56 ± 9.19

The amount of 2-β-hydroxy-1,8-cineole formed WT enzyme and variants KB1 – KB8 is shown in Figure 4.4.5. The variants identified from two iterative rounds of SeSaM exhibited improvement in activity with NADPH as well as with Zinc dust and Cobalt(III)sepulchrate (Figure 4.4.7).
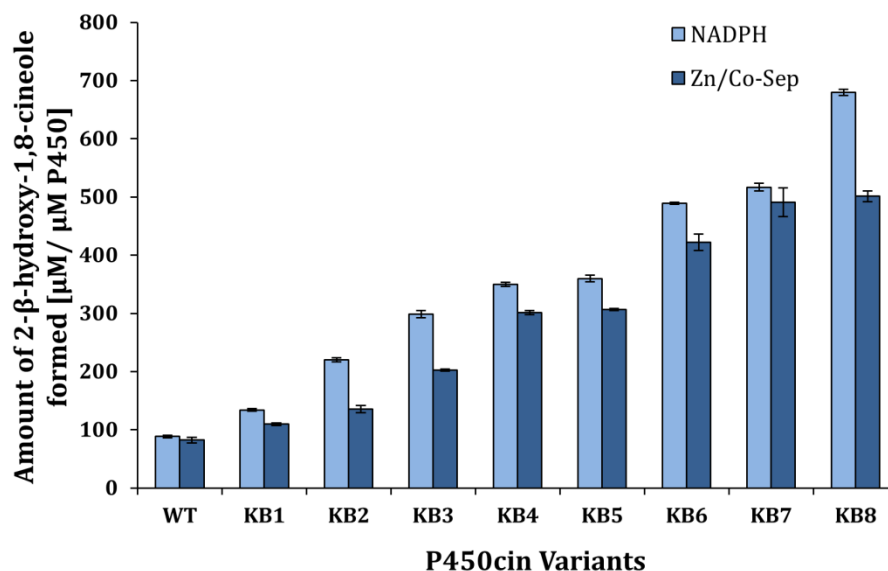
Figure 4.4.7. Amount of 2-β-hydroxy-1,8-cineole for P450cin variants identified after two iterative rounds of SeSaM libraries using NADPH and Zinc dust and Cobalt(III)sepulchrate as electron source. All measurements were performed in triplicate using NADPH [0.4 mM] or using Zinc dust [5 mg] and Cobalt(III)sepulchrate [5 mM] and were performed for 2 h at 30 °C using 1 μM P450. WT (CinA-linker 10aa-CinC)**,** KB1 (Q385H), KB2 (V386S), KB3 (Q385H/V386S), KB4 (Q385H/V386S/T77C), KB5 (Q385H/V386S/L88N), KB6 (Q385H/V386S/T77S/L88C), KB7 (Q385H/V386S/T77C/L88Y) and KB8 (Q385H/V386S/T77C/L88Y).

Thus the improvement in activity for all the identified P450cin variants was observed for both electron transfer systems employed- NADPH as well as Zinc dust and Cobalt(III)sepulchrate. Based on these results, improving activity of enzyme variants for Cobalt(III)sepulchrate as alternative cofactor system instead of NADPH only states increased product formation and improved activity of the CinA-CinC fusion protein towards 1,8-cineole but an assertion about electron transfer rates cannot be postulated. Subsequently, variants KB3 and KB8 were selected for determination of kinetic parameters (k_{cat} and K_m) because they had the highest improvements in reactions performed with Zn/Co^{III}-Sep as alternative cofactor system (see section 4.4.6). Improvement in activity for the identified P450cin variants was also validated using bioelectrocatalytic reaction set up (discussed in Chapter 4, section 4.6).

4.4.6. Characterization of WT and Improved Variants (KB3 and KB8)

Characterization of WT (CinA-linker10aa-CinC) and two improved variants KB3 and KB8 was performed employing electron donor systems Cobalt(III)sepulchrate. Purification of these variants was performed as described in chapter 3; section 3.3.4. In order to determine the range of substrate concentration for determination of V_{max} and K_m, the substrate concentration was changed from 10 µM to 5 mM and enzyme concentration (P450 concentration) of 1 µM (Figure 4.4.8). Kinetic parameter (V_{max} and K_m) were calculated on basis of non-linear regression based on Michaelis -Menten model using GrahPad Prism software.
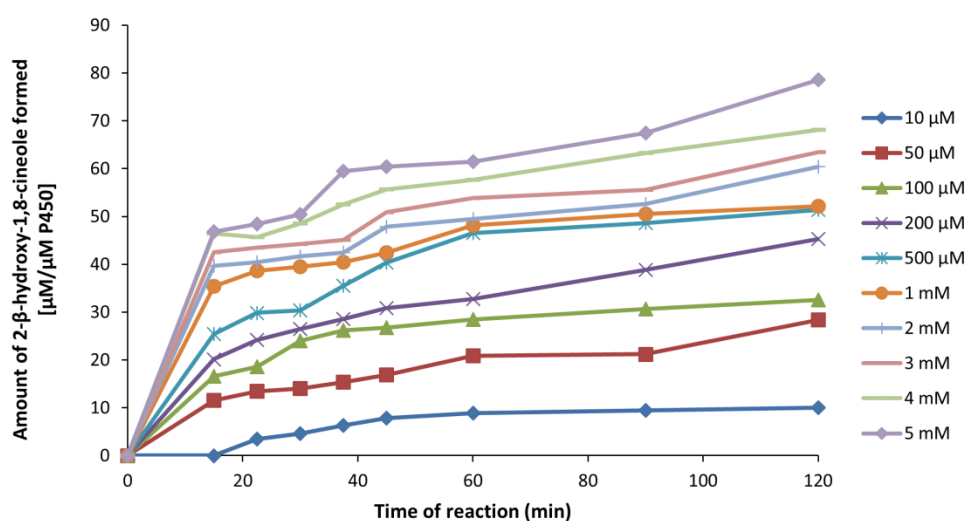
Figure 4.4.8. Activity profile of reaction using 1,8-cineole concentration range from 10 µM to 5 mM for determination of k_{cat} and K_m. Reactions were performed using purified WT enzyme (1 µM; CinA-linker10aa-CinC). Reaction products were extracted with ethyl acetate after 20, 40, 60, 80, 100 and 120 min and analyzed by GC-FID.

Figure 4.4.8 shows the 2-β-hydroxy-1,8-cineole formation profile for WT enzyme using 1,8-cineole concentration range from 10 µM to 5 mM. The concentration of enzyme used for reaction was 1 µM (normalized to P450 concentration). However, using this substrate and enzyme concentration, a linear curve was not obtained (Figure 4.4.8). The reaction reached saturation within 20 min of reaction time. As a result, the concentration of enzyme was decreased to 0.25 µM and the concentration range of 1,8-cineole from 100 µM to 3 mM

(Figure 4.4.9). 1,8-cineole concentration range of 100 μM to 3 mM and enzyme concentration of 0.25 μM showed a linear range for product formation (Figure 4.4.9).
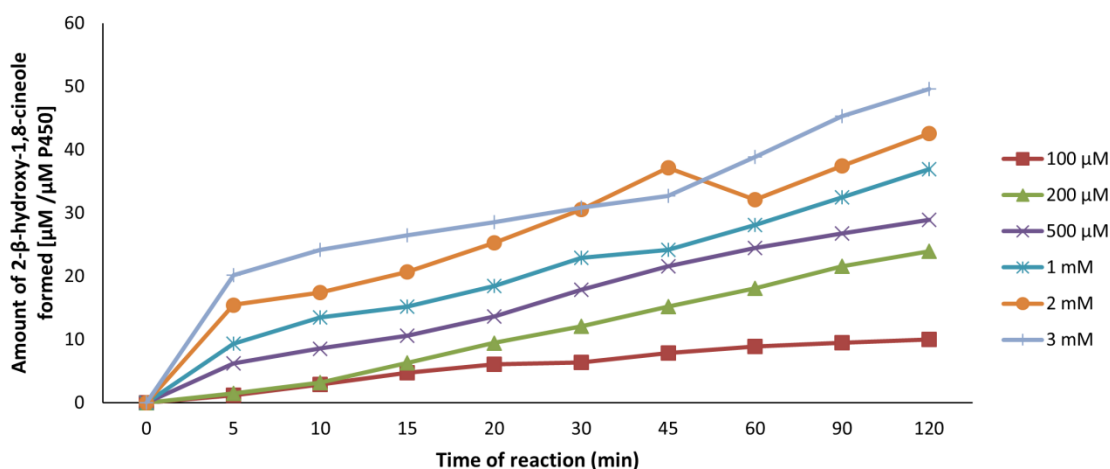

Figure 4.4.9. Activity profile of reaction using 1,8-cineole concentration from 100 μM to 3 mM for determination of k_{cat} and K_m. Reactions were performed using purified WT enzyme (0.25 μM; CinA-linker10aa-CinC). Reaction products were extracted with ethyl acetate after 20, 40, 60, 80, 100 and 120 min and analyzed by GC-FID.

Determination of the kinetic parameters (k_{cat} and K_m) was thus performed with enzyme concentration of 0.25 μM (normalized to P450 concentration) and 1,8-cineole concentration range of 100 μM to 3 mM. Table 4.4.4 summarizes the kinetic parameters for variants KB3 (Q385H/V386S) and KB8 (Q385H/V386S/T77C/L88Y) with improved activity towards 1,8-cineole with electron donor system Cobalt(III)sepulchrate in comparison to wild type (CinA-linker10aa-CinC). Variant KB3 showed a 2 fold improvement in k_{cat} values (1.38 μM min⁻¹) and variant KB8 showed a 3.7 fold increase (2.5 μM min⁻¹) in comparison to the WT (0.67 μM min⁻¹). However the K_m value was unchanged (90 μM) (Table 4.4.5 and Figure 4.4.9).

Table 4.4.5: Kinetic parameters for WT (CinA-linker 10aa-CinC), KB3 (Q385H/V386S) and KB8 (Q385H/V386S/T77C/L88Y). Kinetic parameters (V_{max} and K_m) were obtained by non-linear regression based on the Michaelis- Menten model using GraphPad Prism software package. TTN (total turnover numbers) were obtained after a reaction time of 24 h and were calculated as µM product/µM enzyme. TTN were determined for two different concentrations of 1,8-cineole (1 mM and 5 mM).

Variant	K_{cat} [µM/min]	K_m [mM]	K_{cat}/K_m [µM min^{-1} mM^{-1}]	TTN 1 mM	TTN 5 mM
WT	0.67	0.09	7.4	60 ± 2	260 ± 5
KB3	1.328	0.1	13.28	200 ± 4	562 ± 7.5
KB8	2.5	0.09	27.77	480 ± 2.5	1170 ± 6.5

Another kinetic parameter as a measure of catalytic performance of a P450 system is its total turnover number (TNNs). TTNs were determined using non optimized reaction conditions. Table 4.4.5 gives the total turnover numbers for the WT enzyme and the engineered variants KB3 (Q385H/V386S) and KB8 (Q385H/V386S/T77C/L88Y) which was calculated as µM product µM^{-1} P450 for a reaction time of 24 h employing two different 1,8-cineole concentrations (1 mM and 5 mM). Employing mutants KB3 and KB8 more than 4 fold product concentrations were obtained as compared to the WT.
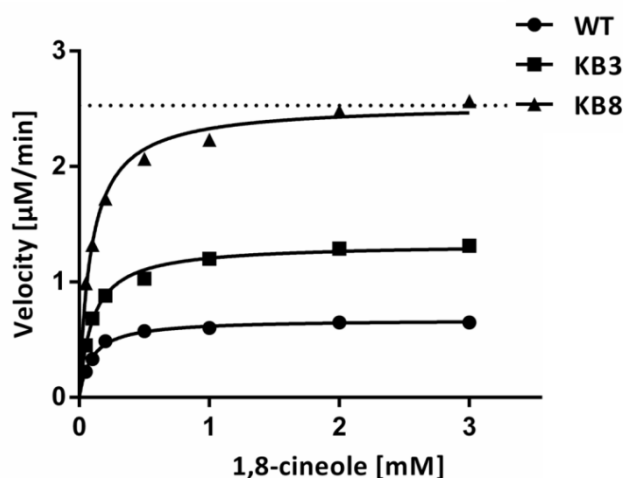
Figure 4.4.10. Kinetic characterization WT (CinA-linker 10 aa-CinC), KB3 (Q385H/V386S) and KB8 (Q385H/V386S/T77C/L88Y) for determination of k_{cat} and K_m with electron source Cobalt(III)sepulchrate.

To summarize the results, values determined for k_{cat} and K_m for P450cin variants performed using purified enzyme confirmed the screening data. Variant KB3 showed a 2 fold improvement in k_{cat} values (1.38 µM min^{-1}) and variant KB8 showed a 3.7 fold increase (2.5 µM min^{-1}) in comparison to the WT (0.67 µM min^{-1}). However the K_m value was unchanged (90 µM) (Table 4.4.5 and Figure 4.4.10).

4.4.7 Discussion

Redox proteins by natural design are not suitable for use with alternative co-factor systems. Redox active sites are often embedded deeply in the protein core with tight regulation over the electron transfer mechanism (Wong and Schwaneberg, 2003). Directed protein evolution allows overcoming this challenge by redesigning and improving of redox proteins for alternative cofactor systems without the knowledge of protein structure or electron transfer mechanism. One of the main hurdle in employing P450 monooxygenases for preparative chemical synthesis using cell free systems is their requirement for expensive cofactors such as NA(D)PH. Of many approaches to circumvent this problem is to use an alternative electron system using a mediator which shuttles electrons between a cheap electron source and active site of the protein.

Figure 4.4.11 summarizes the directed P450cin evolution campaign generating a final variant KB8 carrying four mutations (Q385H/V386S/T77N/L88R) with 3.8 times higher catalytic efficiency (k_{cat}/K_m)(27.77 µM^{-1} min^{-1}) compared to WT enzyme (7.4 µM^{-1} min^{-1}) and an increase in k_{cat} from 0.67 µM min^{-1} for WT to 2.5 µM min^{-1} for mutant KB8.
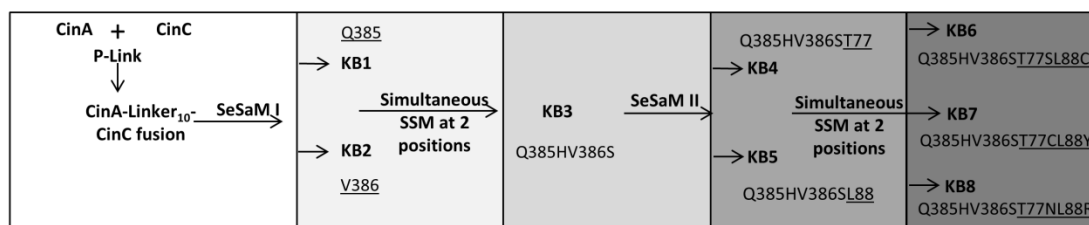

Figure 4.4.11. Summary of the directed P450cin evolution campaign. P450cin was evolved for alternative co-factor system Cobalt(III)sepulchrate to replace NADPH in cell free reactions. Two rounds of directed evolution using SeSaM and Site Saturation Mutagenesis was performed.

In this thesis, electron transfer system Zinc dust with mediator Cobalt(III)sepulchrate was employed as alternative reduction equivalents for conversion of 1,8-cineole to 2-β-hydroxy-1,8-cineole using P450cin fusion protein variants.

Screening of first round of random library was performed using NADPH as reducing equivalent. NADPH was used since the detection limit for TLC was high (100 μM or more) and using Zinc dust and Cobalt(III)sepulchrate about 82 μM μM^{-1} P450 for WT enzyme (CinA-linker 10aa- CinC) was obtained. Consequently, the screening for first round of SeSaM library was performed using NADPH to identify a variant with increased activity for 1,8-cineole which could be further used in subsequent rounds of random mutagenesis for evolution towards Zinc dust and Cobalt(III)sepulchrate. Screening of mutant libraries in subsequent round of SeSaM and site saturation mutagenesis libraries was achieved using Zinc dust and Cobalt(III)sepulchrate. The amino acid positions identified from two iterative rounds of random mutagenesis are shows in Figure 4.4.12.
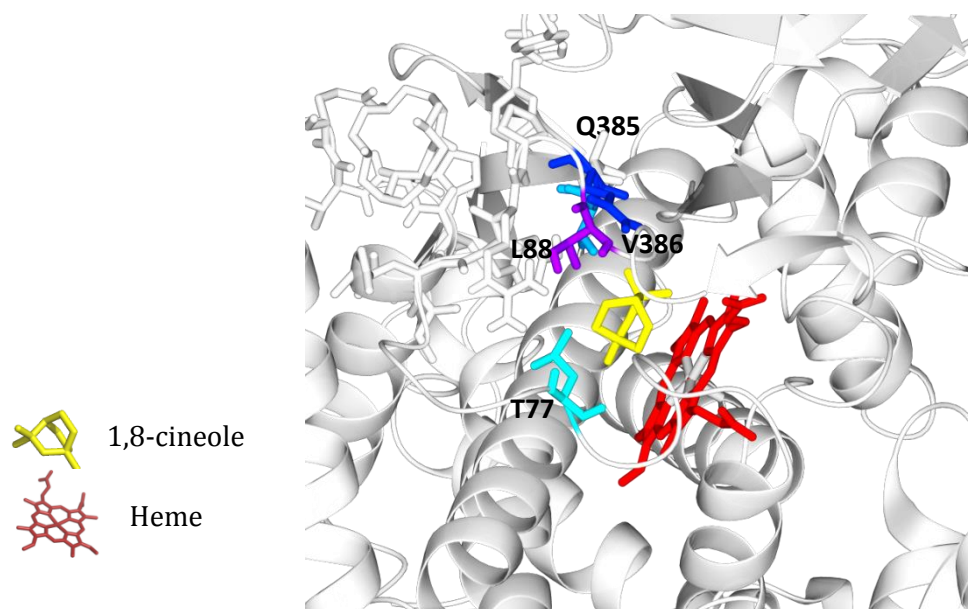
Figure 4.4.12. Crystal structure of P450cin (CinA; PDB id: 1T2B); showing the positions Q385, V386, T77 and L88 identified after two round of random evolution of P450cin fusion.

P450cin closely resembles P450cam not only in size, structure but also the chemical composition of 1,8-cineole and camphor, their respective substrates (Meharenna et al.,

2004). Extensive protein engineering studies have been reported on P450cam enabling the understanding of role of active site residues in cofactor binding and substrate specificity (Fasan, 2012;Urlacher et al., 2004). In substrate-protein contact regions where the secondary structural elements are conserved between P450cin and P450cam there is in general a one-to-one correspondence in the location, but not identity, of residues contacting the substrate. Amino acid residues T77, L88, L237, V386 and A285 constitute the substrate binding pocket for P450cin (Meharenna et al., 2004). Amino acid Val386 which is a part of the nonpolar substrate pocket in P450cin, correspond to Val396 in P450cam. Whereas, residues T77 and L88 in P450cin correspond to amino acid residues F87 and F98 in P450cam (Meharenna et al., 2004). Amino acids residues F87, F98, V396 in P450cam are responsible for access of camphor to the active (Poulos et al., 1987). Due to correspondence in the location of amino acids in P450cam and P450cin it can be speculated that the amino acids identified to be improved for activity for P450cin might influence the access or binding of 1,8-cineole in the active site of the enzyme.

In case of mutant KB3 (Q385H/V386S), amino acids with different chemically properties was substituted only at one the position. The uncharged amino acid residue glutamine was replaced with histidine, a positively charged amino acid. However, property of amino acid at position 385 remained unchanged (valine to serine).

In case of mutant KB8 (Q385H/V386S/T77N/L88R), the non-polar aliphatic residue at position 88 (leucine) was substituted with polar positively charged arginine whereas the property of the amino acid remained unchanged at position 77 (threonine to asparagine).

The binding of the mediator Cobalt(III)sepulchrate in the protein matrix or the amino acid residues crucial for electron transfer in cytochrome P450cin have not been elucidated yet. Two possible hypotheses can be stated for the interaction between the mediator and the protein, one in which the mediator does not enter the protein matrix but transfers electrons to the electron transfer chain in the protein matrix (Nazor, 2007; Ruff, 2012) like as reported in P450 BM3. Another possible hypothesis could be that the mediator binds in close proximity to the electron transfer chain or near the heme-FMN interface and the electron transfer proceeds through electrostatic interaction between the mediator and protein matrix. A molecular understanding on how amino acids substitutions identified in this report contribute to increase in activity with Zinc dust and Cobalt(III)sepulchrate is

however challenging and extensive computational as well as crystallographic studies might provide insights regarding the same would further allow to verify the obtained experimental data.

The optimal ratio reported for P450cin using reconstituted system consisting of CinA, *E. coli* Fpr and CinC is 1:2:8 (Hawkes et al., 2010). Furthermore they reported a catalytic efficiency (k_{cat}/K_m) of 21 μM min $^{-1}$ μM $^{-1}$ P450 with a ratio of 1:1:1 (CinA: Fpr: CinC) using NADPH as electron source. Using P450cin variant KB8 (CinA:CinC ratio 1:1), a comparable catalytic efficiency of 27.77 μM min $^{-1}$ μM $^{-1}$ P450 with electron transfer system Zinc dust and Cobalt(III)sepulchrate was obtained (Table 4.4.5). Consequently mutant KB8 in combination with Zinc dust and Cobalt(III)sepulchrate can efficiently replace the reconstituted P450cin system and NADPH.

Slessor et al., reported construction of vectors for the polycistronic co-expression of P450cin and cindoxin (CinC) with or without terminal redox partner (*E. coli* Fpr). However, when Fpr was coexpressed with P450cin and cindoxin, the expression levels of cindoxin were drastically decreased presumably due to competition between cindoxin and Fpr for flavin cofactors. The co-expression system also decreased the expression levels of P450cin (Slessor et al., 2012). Whereas the fusion protein reported in this thesis posed no difficulties in expression levels, neither for CinA or CinC (see Chapter 3, Section 3.3.4). Use of fusion protein not only facilitates the expression but also simplifies P450cin system on genetic level making it suitable for directed evolution experiments. The use of Zinc dust and Cobalt(III)sepulchrate in combination with CinA-CinC fusion protein not only circumvents the use NADPH but also of CinB, a non-expressible natural redox partner of P450cin.

The identified P450cin variants for improvement in activity with Zinc and Cobalt(III)sepulchrate were further validated with electrobiocatalytic reaction set up. The results are presented and discussed in Chapter 4, Section 4.6.

4.4.8 Conclusion

In order to improve the fusion of P450cin protein (CinA-linker 10 aa-CinC) for mediated electron transfer system (Zinc dust and Cobalt(III)sepulchrate) two random mutagenesis libraries were generated with the SeSaM method. Screening of the SeSaM libraries performed using a validated TLC based screening system determined four positions Q385,

V386, T77 and L88 to be crucial for improvement in activity towards 1,8-cineole. Variant KB8 (Q385H/V386S/T77N/L88R) (k_{eff} = 27.77 μM min^{-1} mM^{-1}) exhibited having a catalytic efficiency (k_{cat}/K_m) 3.8 times as the starting variant (CinA-linker10aa-CinC) (k_{eff} = 7.4 μM min^{-1} mM^{-1}). Variant KB3 (Q385H/V386S) showed nearly 2 fold improvement in k_{cat} values (1.38 μM min^{-1}) and variant KB8 showed a 3.7 fold increase (2.5 μM min^{-1}) in comparison to the WT (0.67 μM min^{-1}) with no significant change in the K_m value (90 μM). Additionally the total turnover numbers for the variants KB8 were determined to be 1170 and 480 under non optimized conditions employing Zinc dust and Cobalt(III)sepulchrate and 5 mM and 1 mM 1,8-cineole concentration respectively. Variant KB8 showed a 1.5 times (electron source: Zinc dust and Cobalt(III)sepulchrate) improvement compared to optimal equimolar mixture of CinA, CinC and Fpr (electron source: NADPH). The fusion of CinA and CinC has a ratio of 1:1 compared to the optimal ratio of 1:2:8 for CinA: Fpr: CinC. Furthermore, CinA-CinC fusion protein in combination with Zinc and Cobalt (III) sepulchrate omits the use of CinB, the natural redox partner of P450cin which is yet to be expressed in an active form.

4.5 Semi-Rational Design of Cytochrome P450cin Fusion Protein for Mediated Electron Transfer

4.5.1. Aim of the Work

Semi-rational approach has been used extensively in directed protein evolution campaigns (Kardashliev et al., 2013; Martinez et al., 2013) for improving enzyme properties. Semi-rational approaches target, specific amino residue(s) to mutate on the basis of prior structural or functional information to create 'smart' libraries with reduced screening efforts that are more likely to yield positive results (Chica et al., 2005). This chapter focuses on the semi-rational approach used for engineering of P450cin fusion protein for mediated electron transfer.

4.5.2 Homology Modeling for P450cin Fusion Protein

A homology model of CinA-linker10aa-CinC was generated using homology modeling routine of the YASARA software (Krieger et al., 2002). Homology model of CinA-linker10aa-CinC was generated as described in Chapter 4, Section 4.3.2. P450cin was aligned with the crystal structure of P450 BM3 in complex with cobalt sepulchrate (Shehzad et al., 2013) (unpublished results). Although, the similarity between P450 BM3 and P450cin is < 20%; P450 BM3 was selected for structural alignment since both P450cin and P450 BM3 have FAD and FMN containing redox partners. Crystal structure of P450cin was aligned with crystal structure of P450 BM3 in complex with cobalt-sepulchrate both based in a full sequence alignment and a partial aligment based in the loop region of P450 BM3 (residues 187 to 198) (residues 175 to 179 in P450cin) in order to estimate the position of simulation box for docking for Cobalt(III)sepulchrate. The charge of cobalt ion was fixed to +3 and the sepulchrate cage was placed around the cobalt ion. A total of 100 docking runs were performed and the highest ranked docking poses were selected for further analysis. (Figure 4.5.1 A and B).
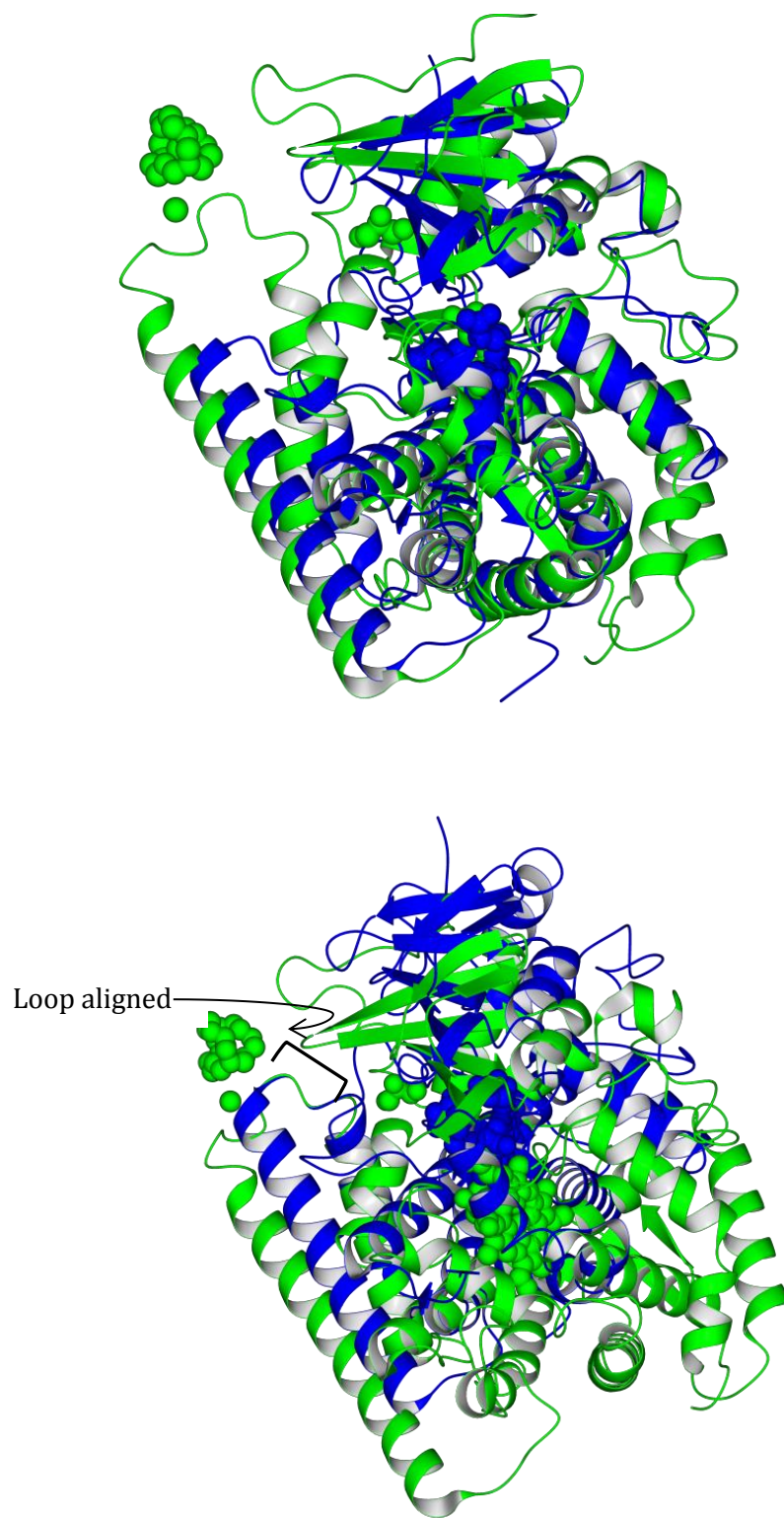
Figure 4.5.1. Structural alignment of P450 BM3 (green) and P450cin homology model (blue). **(A)** Helix of P450cin was aligned with helix of P450 BM3. **(B)** Loop of P450cin was aligned with loop of P450 BM3.

4.5.3 Docking of Cobalt(III)sepulchrate in P450cin Fusion Protein

Cobalt(III)sepulchrate was docked in the constructed homology model using YASARA on the basis of the crystal structure of P450 BM3 in complex with cobalt(III)sepulchrate as reported by Shehzad et al., 2013 (unpublished results). The charge of cobalt ion was fixed to +3 and the sepulchrate cage was placed around the cobalt ion. YASARA uses a similar method for docking of ligands in protein matrix like Autodock (Goodsell et al., 1996) with certain additional features. It allows interactively placing the simulation cell around the active site to focus docking on the most important region and to perform rigid as well as flexible docking. Simulation cell defines the space and position in the protein structure where a target substrate or molecule will be docked. Two approaches were used to determine the position of simulation cell in the P450cin structure for docking of Cobalt(III) sepulchrate using YASARA. In the first approach, structural alignment of CinA-linker10aa-CinC and P450 BM3 was used. P450 BM3 contained the Cobalt(III)sepulchrate docked to it. A simulation cell was placed around the Cobalt(III)sepulchrate and the structure of P450 BM3 was deleted leaving the homology model of CinA-linker10aa-CinC with a simulation cell. Cobalt(III)sepulchrate was subsequently docked in this simulation cell using YASARA docking routine.
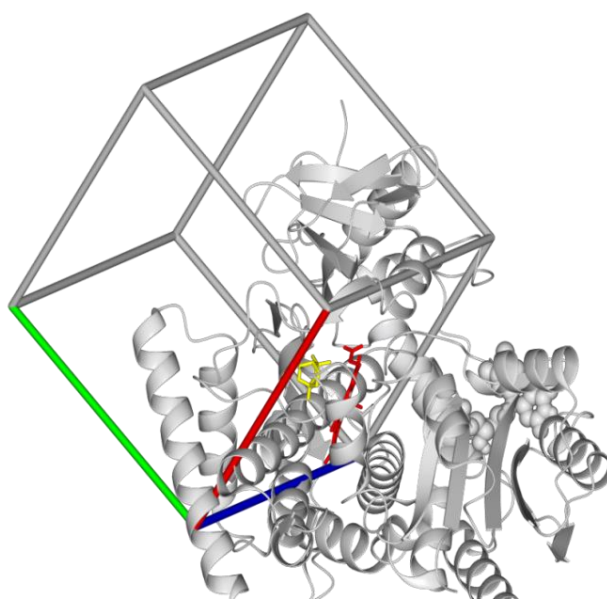
Figure 4.5.2. Homology model of CinA-linker10aa-CinC with the simulation cell for docking of Cobalt(III)sepulchrate. The box represents the simulation cell. 1,8-cineole is shown in yellow and heme is shown in red.

In the second approach, the homology model constructed for CinA-linker10aa-CinC was used and the simulation cell was set up in the active site of the enzyme. Using this, docking of Cobalt(III)sepulchrate was performed. In total 100 docking runs were performed using both the approaches. The results obtained from both the approaches were similar and the position of Cobalt(III)sepulchrate was exactly in the same position. Figure 4.5.2 shows a homology model of CinA-linker10aa-CinC and the position of the simulation cell.

Another set of homology models for docking of Cobalt(III)sepulchrate were prepared in which only the crystal structure of CinA (PDB id: 1T2B) was used, since all the mutations identified from random mutagenesis library were found in the heme domain (CinA). The docking of Cobalt(III)sepulchrate in the crystal structure of CinA was performed using YASARA in a similar way as described above for CinA-CinC fusion protein with linker length of 10 amino acids. The model generated after docking of Cobalt(III)sepulchrate for CinA was compared to the model generated for CinA-CinC fusion protein with linker length of 10 amino acids. However, no significant changes were observed in both the generated models.

4.5.4 Selection of Amino Acid Residues

Docking results were analyzed taking into consideration the first five docking poses of Cobalt(III)sepulchrate. The docking poses are ranked according to the most stable conformation to the least stable configuration (lowest energy to highest energy). For instance pose 1 will have more stability and less energy than pose 2 and hence will be favored. The results of docking for Cobalt(III)sepulchrate were analyzed for first five 'best' poses. Figure 4.5.3 shows five different poses of Cobalt(III)sepulchrate (represented by shades of purple).

Amino acid residues identified from random muatagenesis (Q385, V386, T77, and L88) along with the proposed docked Cobalt(III)sepulchrate pose in P450cin are shown in Figure 4.5.3 and Figure 4.5.4. Upon evaluation of the proposed docking site for Cobalt(III)sepulchrate, residues 79, 80 and 81 were selected as they were in close proximity from both the amino acid positions identified by random mutagenesis (77 and 88) and in direct contact with Cobalt(III)sepulchrate. In this way, residues 79, 80 and 81 could complete a putative 'path' between Cobalt(III)sepulchrate and the active site.
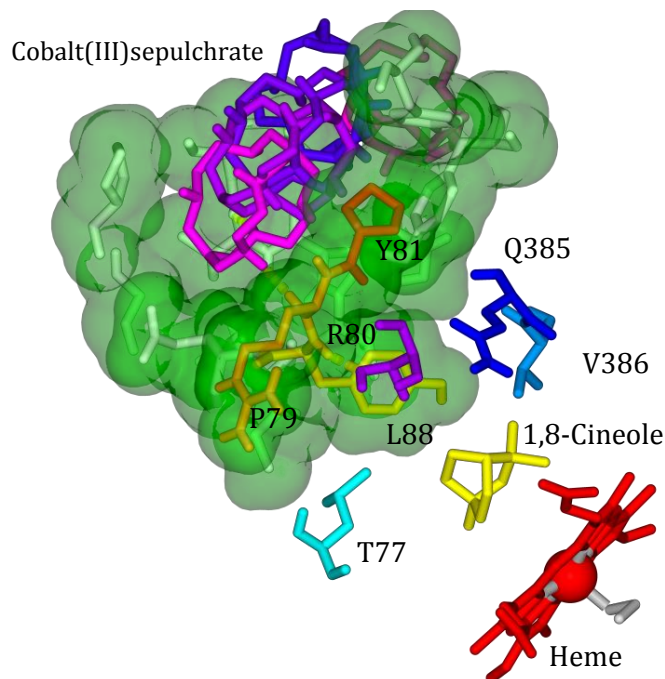
Figure 4.5.3. Five best docking poses for Cobalt(III) sepulchrate along with positions identified from random mutagenesis (Q385, V386, T77 and L88). The figure also shows the position selected for saturation by semi-rational approach (amino acid positions 79, 80 and 81).
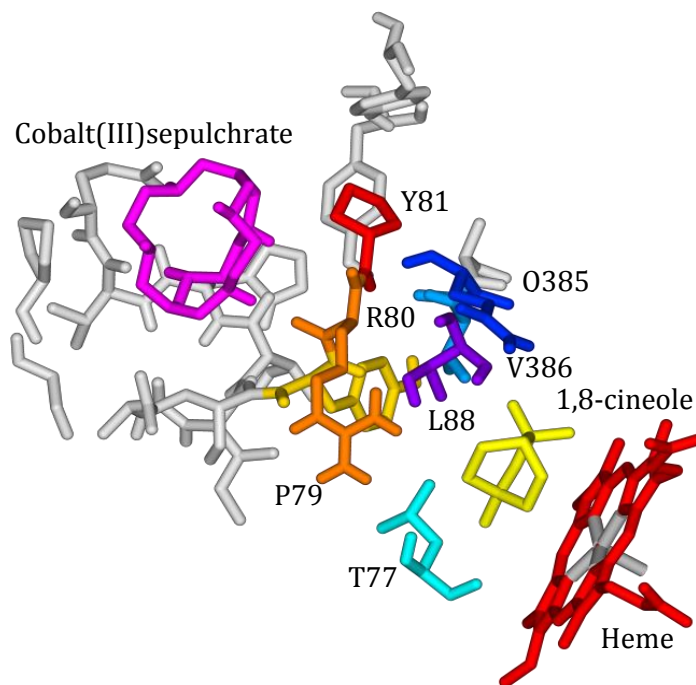

Figure 4.5.4. Best docking pose for Co (III) sepulchrate along with positions identified from random mutagenesis (Q385, V386, T77 and L88). The figure also shows the position selected for saturation by semi-rational approach (amino acid positions 79, 80 and 81).

The amino acid residues 79, 80 and 81 were selected on assumption that the binding of Cobalt(III)sepulchrate would occur at a position similar to that observed in the docking model (Figure 4.5.4). This would in turn facilitate the transfer of electrons from Zinc via Cobalt(III)sepulchrate through the selected amino acid residues (79, 80 and 81) towards heme for conversion of 1,8-cineole to 2-β-hydroxy-1,8-cineole.

4.5.5 Generation of Simultaneous Site Saturation Mutagenesis Library at Position 79, 80, and 81

The positions selected (79, 80, and 81) after docking of Cobalt(III)sepulchrate were simultaneously saturated using a two-step protocol as described in chapter 3, section 3.2.5. The primers used contained NNK codon degeneracy and are listed in table 5; appendix section. Variant KB3 (Q385H/V386S) was used as template for generation of simultaneous SSM library at position 79, 80, and 81. Mutant KB3 was obtained after simultaneous site saturation at positions Q385 and V386 identified after SeSaM round I. The amplification of target gene was confirmed on a 0.8% agarose gel which showed a band at the expected size of 3.8 kbp (Figure 4.5.5). The PCR product was digested with *DpnI* and transformed in chemically competent *E. coli* BL- 21 Gold (DE3) lacI[Q1] cells (Chapter 3, Section 3.1.4). Three random clones were sequenced before library was screened to identify an improved variant with alternative cofactor system. The sequencing result showed randomization of all three amino acid positions (Table 4.5.1).
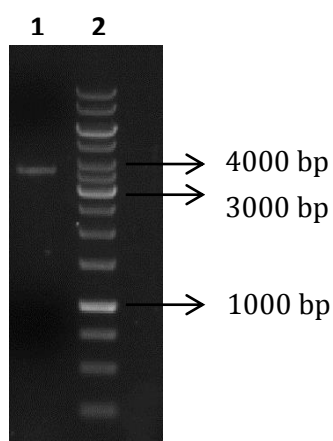
Figure 4.5.5. 0.8% agarose gel for simultaneous site saturation mutagenesis at position 79, 80 and 81. Lane 1 = SSM at position 79, 80 and 81 (expected size ~ 3.8 kbp). Lane 2 = 1 kbp DNA ladder.

Table 4.5.1. Sequence analysis for three random clones from the simultaneous site saturation mutagenesis library at position 79, 80, and 81 (WT = wild type enzyme).

	Position		
Clone no	**79**	**80**	**81**
1	cat	atg	agt
2	ccg	cat	gag
3	tct	cgt	ttt
WT	ccg	cgt	tat

4.5.6 Screening of Simultaneous Site Saturation Mutagenesis Libraries

Screening of 450 clones from the simultaneous SSM library was performed using thin layer chromatography using crude cell lysates in a microtiter plate using electron source Cobalt(III)sepulchrate (described in chapter 3; section 3.3.2). The clones with hydroxylation activity higher than 20% of WT (Q385H/V386S) towards 1,8-cineole were selected for rescreening. For Rescreening of 30 clones was perfomed by expressing the selected variants in 200 mL TB media (described in chapter 3, section 3.2.8) and reactions were performed using known amount of P450 (1 µM) as described in chapter 3; section 3.3.2 and analyzed by GC-FID (chapter 3; section 3.4.3).
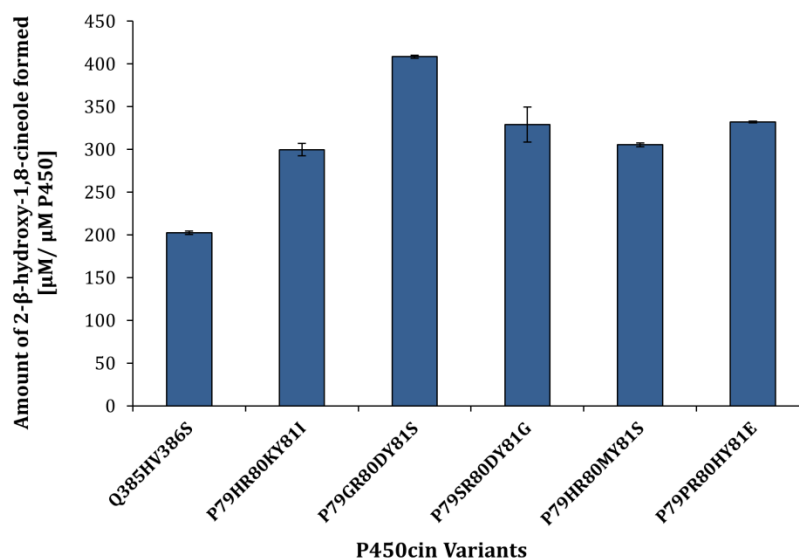
Figure 4.5.6. 2-β-hydroxy-1,8-cineole formation for five best mutants identified from simultaneous saturation library at position 79, 80, and 81.

Only 450 clones from the simultaneous saturation library were screened owing to the low throughput screening capacity using the TLC based detection (~200 clones /day) covering about < 0.0001% fraction of the theoretical sequence space (Reetz et al., 2008). Figure 4.5.6 shows the hydroxylation activity of five best variants identified after screening of SSM library at position 79, 80, and 81. Mutant KB9 (Q385H/V386S/P79G/R80D/Y81S) was identified which showed approximately 1.5 times activity as that of starting variant KB3 (Q385H/V386S) (Figure 4.5.6) for reactions using cell free lysates.

In order to study the cooperative effect of combining the "key beneficial" mutation identified by random and rational evolution; amino acid positions of identified variants KB8 (Q385H/V386S/T77N/L88R) and KB9 (P79G/R80D/Y81S) were combined by site directed mutagenesis to generate mutant KB10(Q385H/V386S/T77N/L88R/P79G/R80D/Y81S). However, interestingly the mutant showed no hydroxylation activity (Table 4.5.2).

Table 4.5.2. Activity profile for variants KB8, KB9 and KB10. Reactions were performed for 2 h at 30 °C using 1 µM P450; Zinc dust [5 mg] and Cobalt(III)sepulchrate [5 mM] in triplicate. Reaction products were extracted with ethyl acetate and analyzed by GC-FID for determination of 2-β-hydroxy- 1,8-cineole.

Variant	2-β-hydroxy-1,8-cineole with Zinc dust and Cobalt (III) Sepulchrate (µM µM P50^{-1})
KB8 (Q385H/V386S/T77N/L88R)	501.56 ± 9.15
KB9 (Q385H/V386S/P79G/R80D/Y81S)	408.13±1.8
KB10 (Q385H/V386S/T77N/L88R /P79G/R80D/Y81S)	inactive

4.5.7. Discussion

Semi-rational approach has been used extensively in directed protein evolution campaigns (Kardashliev et al., 2013; Martinez et al., 2013) for improving enzyme properties. Semi-rational approach targets, specific amino residue(s) on the basis of prior structural or functional information to create 'smart' libraries that are more likely to yield positive results (Chica et al., 2005). Using homology modelling, amino acid residues (79, 80 and 81) which could have an impact on the activity of the fusion protein were identified (Figure 4.5.3). These amino acids were in close proximity to the docked Cobalt(III)sepulchrate

which would in turn facilitate the transfer of electrons from Zinc via Cobalt(III)sepulchrate through the selected amino acid residues (79, 80, and 81) towards heme for conversion of 1,8-cineole to 2-β-hydroxy-1,8-cineole. The chemical properties of amino acid residues identified from the SSM library were drastically different. In case of the best mutant KB9 (Q385H/V386S/P79G/R80D/Y81S), cyclic proline was replaced to an aliphatic glycine, acidic aspartic acid replaced the basic arginine and aromatic tyrosine was substituted by non-aromatic serine. In other identified mutants a similar pattern of substitution was observed. In mutant Q385H/V386S/P79H/R80K/Y81I, cyclic proline was substituted with basic positively charged histidine, arginine with lysine, amino acids with similar properties and aromatic tyrosine with an aliphatic isoleucine. The mutant Q385H/V386S/P79S/R80D/Y81G showed a substitution from cyclic to non-aromatic at position 79, basic amino acid substituted with acidic amino acid at position 80 and an aromatic tyrosine substituted with aliphatic glycine.

However no particular amino acid or amino acids with similar properties was found to be favored at a particular amino acid position. As a result it is difficult draw a decisive conclusion regarding the role and effect of each amino acid substitution at positions 79, 80, and 81. A further look at any identified amino acid substitution was performed using YASARA where models for protein with amino acid substitution identified at position 385, 386, 77, 88, 79, 80 and 81 were generated. The protein models were generated with an aim to determine if the amino acid substitutions influence any structural or conformational changes in the active site or protein structure. The generated models were then compared to the wild type enzyme in order to observe if the substitutions led to a drastic effect on the orientation, conformation or structure of the active site of the enzyme. However, no major structural changes or formation or disruption of hydrogen bonds, were observed. The binding of the mediator Cobalt(III)sepulchrate in the protein matrix or the amino acid residues crucial for electron transfer in cytochrome P450cin have not been elucidated yet. A molecular understanding on how amino acids substitutions identified in this report contribute to increase in activity with Zinc and Cobalt(III)sepulchrate is however challenging and extensive computational as well as crystallographic studies might provide insights regarding the same would further allow to verify the obtained experimental data.

4.5.8. Conclusion

Semi-rational approach for selection of amino acid residues was based on the docking of Cobalt(III)sepulchrate in P450cin fusion protein. Amino acid residues 79, 80 and 81 were selected as they were in close proximity from the amino acid positions identified by random mutagenesis (77 and 88) and were in direct contact with Cobalt(III)sepulchrate. In this way, residues 79, 80 and 81 could complete a putative 'path' between Cobalt(III)sepulchrate and heme in the active site of the enzyme. Screening of simultaneous site saturation library (position 79, 80, and 81) identified a variant KB9 (Q385H/V386S/P79G/R80D/Y81S) with about 1.5 times increased hydroxylation activity (408.13±1.8 µM µM P450^{-1} of 2-β-hydroxy-1,8-cineole) compared to variant KB3 (Q385H/V386S) (501.56 ± 9.19 µM µM P50^{-1} of 2-β-hydroxy-1,8-cineole).

Though the improvement is not significant, semi-rational approach of P450cin however demonstrates that it is possible to improve catalytic activity of P450cin towards 1,8-cineole. Further investigations focusing on combined mutations and molecular modeling attempts will help to reveal structure-function relationships for a molecular understanding and role of identified amino acid positions in P450cin.

4.6 Electrochemical Investigation of Identified P450cin Variants[‡‡]

4.6.1 Aim of the Work

Electrochemical approaches for replacement of NAD(P)H have been described for cytochrome P450cam, cytochrome P450 BM3. This part of the thesis focuses on electrochemical investigations of P450cin variants identified to be improved in solution with Zinc dust and Cobalt(III)sepulchrate.

4.6.2 Electrochemical Set Up

The electrochemical reaction set up used for performing electrobiocatalytic reactions for the identified P450cin variants is as described in Chapter 3, Section 3.3.2. All the experiments described in this chapter were performed according to an optimized protocol established at Dechema Research Institute, Frankfurt, Germany. The concentration of Cobalt(III)sepulchrate, potential applied, concentration of enzyme used in the reaction was as described previously (Çekiç et al., 2010).

Cobalt(III)sepulchrate was selected as the mediator since its redox potential (-615 mV vs Ag/AgCl) was more negative than the redox potential of the enzymes active site (P450cin- -402 mV vs Ag/AgCl and cindoxin— -430 mV vs Ag/AgCl) (Andersson et al., 1997; Kimmich et al., 2007). The concentration of mediator for performing the reaction was selected as 5 mM and the potential applied was -750 mV as reported to be optimum for P450cin catalyzed electrobiocatalytic conversion of 1,8-cineole (Çekiç et al., 2010). The reaction set up used for electrobiocatalytic mediated reaction using P450cin is as shown in Figure 4.6.1.

[‡‡] The work was performed in collabration with Dr. Dirk Holtmann at Dechema Research Institute, Frankfurt, Germany.
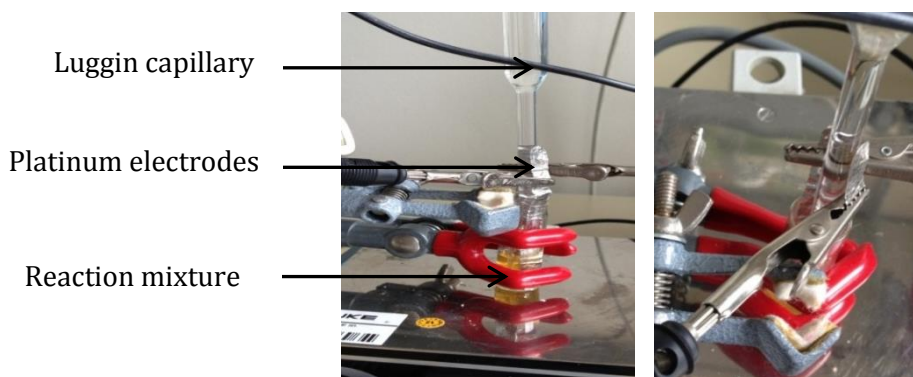
Figure 4.6.1. Electrochemical reaction set up for P450cin mediated conversion of 1,8-cineole to 2-β-hydroxy-1,8-cineole.

4.6.3 Effect of Luggin Capillary Diameter

The luggin capillary was available in two different sizes with internal diameter of 4 mm and 6 mm. The effect of the internal diameter of luggin capillary on amount of 2-β-hydroxy-1,8-cineole formation was determined. The luggin capillary is used to control the placement of the reference electrode relative to the working electrode. The luggin capillary holds an electrolyte, in this case, 0.5 M sodium sulphate.
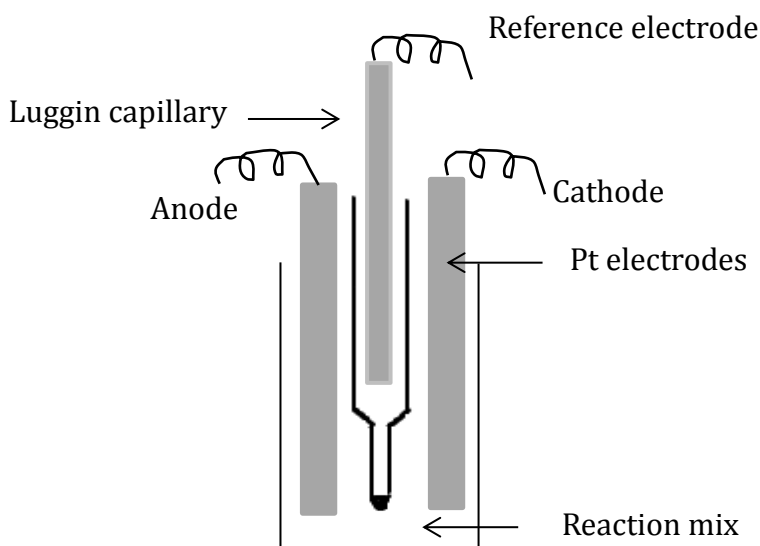
Figure 4.6.2. Electrochemical reaction set up showing the arrangement of electrodes (reference, cathode, and anode), luggin capillary and reaction mixture.

The luggin tip is significantly smaller than the reference electrode itself. It allows sensing of the solution potential close to the working electrode without the adverse effects that occur when the large reference electrode is placed near the working electrode. The resistance of this electrolyte adds to the reference electrode impedance. Larger diameter, shorter Luggin capillaries have lower impedance than narrow bore, longer capillaries.
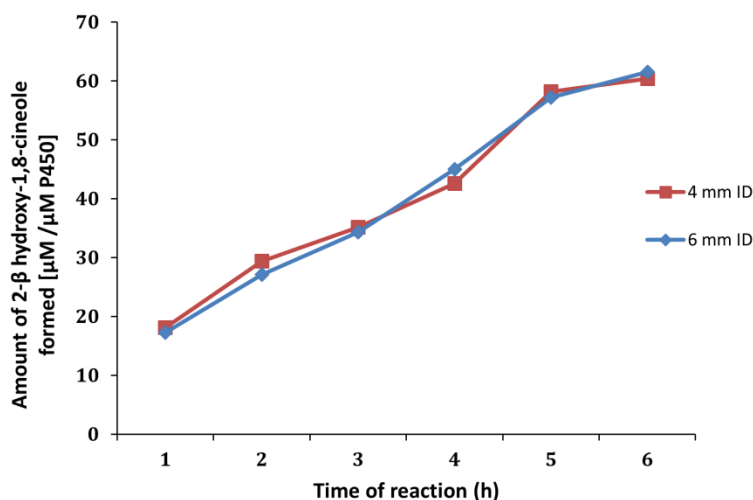
Figure 4.6.3. Influence of luggin capillary internal diameter (4mm and 6 mm) on formation of 2-β-hydroxy-1,8-cineole. The amount of 2-β-hydroxy-1,8-cineole formed was determined using GC-FID for variant KB8 (Q385H/V386S/T77N/L88R) over 6 h of reaction time using 1 µM P450, 5 mM Cobalt(III)sepulchrate and platinum electrodes polarized to -750 mV.

The internal diameter of the luggin capillary did not have significant influence on the formation of 2-β-hydroxy-1,8-cineole (Figure 4.6.3). Formation of 2- β-hydroxy-1,8-cineole was fairly constant using luggin capillary with different diameter. Since, the internal diameter did not influence the product formation, luggin capillary with an internal diameter of 6 mm was chosen for further experiments.

4.6.4 Effect of Electrode Surface Area

In order to determine the effect of electrode surface area on product formation the surface area of electrode was increased from 1 and 2 cm^2. A proportional increase was observed in the amount of 2-β-hydroxy-1,8-cineole formation (Figure 4.6.4). This results show that mediator driven bioelectrocatalysis with P450cin is most probably limited by electrode

surface area. Mediator reduction rate could also become easily the limiting step in the whole reaction and mediator reduction can be limited by electrode surface.
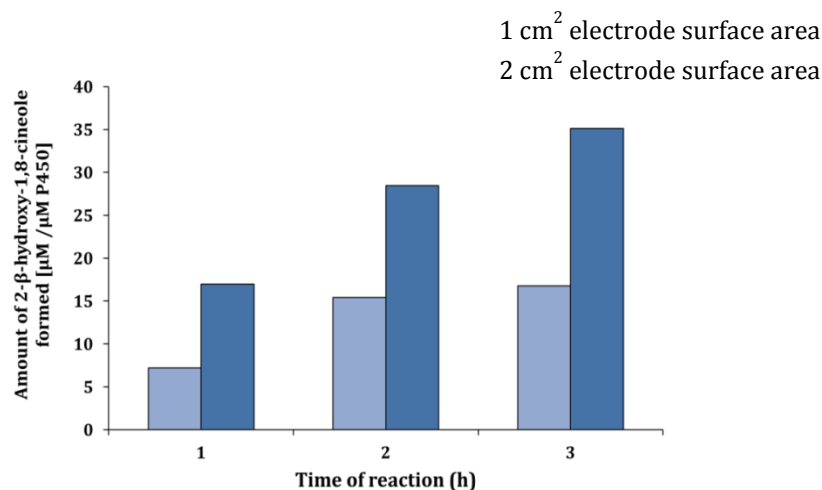
Figure 4.6.4. Effect of electrode surface area on the formation of 2-β-hydroxy-1,8-cineole. The surface area of the electrode was altered from 1cm² to 2 cm². The amount of 2-β-hydroxy-1,8-cineole formed was determined using GC-FID for variant KB8 (Q385H/V386S/T77N/L88R) over 6 h of reaction time using 1 μM P450, 5 mM Cobalt(III)sepulchrate and platinum electrodes polarized to -750 mV.

4.6.5 Stability of CinA and CinC in Electrochemical Reaction

Enzymes upon interaction with the electrode surface usually loose activity. The instability of the enzyme on the surface of the electrode can be attributed to the lower conversion of substrates to products (Holtmann and Schrader, 2007). Thus it was essential to study and monitor the effect of electrochemical reaction set up used in the study on the stability and activity of the P450cin fusion enzyme mutants.

Stability of CinA and CinC in the fusion protein under the applied electrochemical conditions was determined. For this purpose, the two variants Heme (-) and FMN (-) were used (See Chapter 4, section 4.3 for details). The mutant heme (-) had a mutation A91F in heme or CinA domain rendering the heme domain inactive. The FMN (-) mutant had a mutation G594D in the FMN or CinC domain making the FMN domain inactive.

Crude cell lysate for Heme (-) mutant was used to set up a reaction in 50 mM potassium phosphate buffer (pH 7.4) and 1500 U/ml catalase in an electrochemical cell where the working electrode was polarized to −750 mV vs Ag/AgCl for 24 h. At time intervals of 1, 2, 3,

4, 8, 12 and 24 h aliquots of samples were taken and mixed with CinA, *E. coli* Fpr, 1,8-cineole and NADPH and the amount of 2-β-hydroxy-1,8-cineole formed was analyzed by GC-FID. Formation of 2-β-hydroxy-1,8-cineole could be detected for a period of about 5 h. after this time interval no product was detected on GC-FID showing that the activity of CinC was lost almost completely after 5 h of reaction time.

Similarly to determine the stability of CinA, crude cell lysate of FMN (-) mutant was used to set up a reaction in an electrochemical cell polarized to -750 mV vs Ag/AgCl electrode. At time intervals of 1 , 2 , 3 4 ,8 , 12 and 24 h aliquots of samples were taken and mixed with CinC, *E. coli* Fpr, 1,8-cineole and NADPH and the amount of 2-β-hydroxy-1,8-cineole formed was analyzed by GC-FID. Formation of 2- β-hydroxy-1,8-cineole could be detected upto 8 h of reaction time.

4.6.6 Effect of External Addition of CinA and CinC in the Electrochemical Reaction

The results obtained above showed that stability of CinC and CinA in the electrochemical cell was lost after a reaction time of 5 h and 8 h respectively. Previously it has been reported for various enzymes systems that the enzyme loses its stability and hence activity on an electrode surface. In order to check if external addition of CinA and CinC will affect the product formation the following reactions were performed using mutant KB8 (Q385H/V386S/T77N/L88R).

The electrochemical reaction was set up as described in chapter 3, section 3.3.2. In one set of experiment, 1μM, 2 μM and 3 μM of crude cell lysate of CinC was added additionally to the electrochemical reaction. The amount of 2-β-hydroxy-1,8-cineole formed slightly increased when 1 μM of P450cin was added in the reaction. However, addition of 2 μM and 3 μM of P450cin did not significantly increase the 2- β-hydroxy-1,8-cineole formation (Figure 4.6.5 A). In another set of experiment, 1 μM, 2 μM and 3 μM of crude cell lysate of CinC was added additionally to the electrochemical reaction. A similar pattern was observed for CinC as that observed when P450cin was externally added. The amount of 2-β-hydroxy-1,8-cineole formed increased when 1 μM of CinC was added in the reaction. However the product amount did not double or triple for addition of 2 μM and 3 μM CinC (Figure 4.6.5 B).
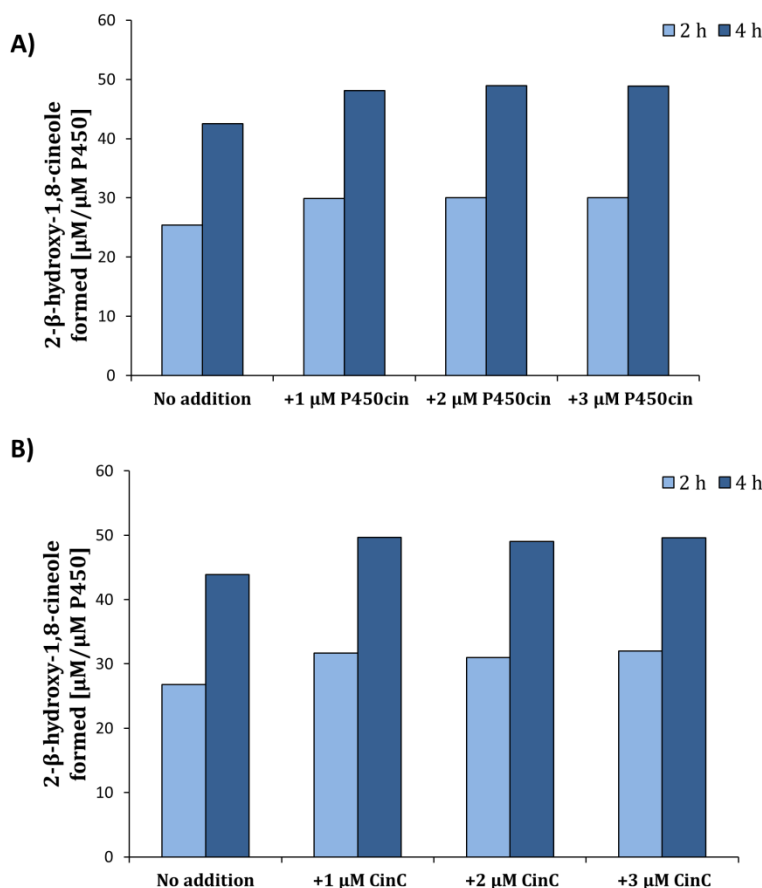
Figure 4.6.5. Effect of external addition of P450cin and CinC in electrobiocatalytic reaction system. **A)** Amount of 2-β-hydroxy-1,8-cineole formed in 2 h and 4 h when 1μM, 2μM and 3μM P450cin was added to the reaction mixture in comparison to when no P450cin was added. **B)** Amount of 2-β-hydroxy-1,8-cineole formed in 2 h and 4 h when 1μM, 2μM and 3μM CinC was added to the reaction mixture. All reactions were performed under potentiostatic control at-750 mV vs Ag/AgCl for mutant KB8 (Q385H/V386S/T77N/L88R).

These results suggest that the CinA and CinC domains in the fusion protein might not be optimally orientated with respect to each other for use in electrochemical reaction setup. However the stability and orientation of protein are not the only factors which can contribute to the productivity of an enzyme system in an electrobiocatalytic reaction. There are numerous other factors to be considered like availability of oxygen, potential applied, surface area of electrode, time of reaction which significantly contributes to the formation of product in an electrobiocatalytic reaction system.

4.6.7. Electrochemical Reaction for Identified P450cin Variants

Using the optimized electrochemical reaction conditions 2-β-hydroxy-1,8-cineole formation was investigated for WT and improved variants (KB1- KB9). The reactions were performed using Platinum electrodes as electron source and Cobalt(III)sepulchrate as mediator as described in chapter 3, section 3.3.2.

The mutants identified to be improved with Zinc dust and Cobalt(III)sepulchrate in solution (KB1- KB9) were tested using electrochemical cell. The aim of the study was to substitute NADPH in P450cin catalysed reaction systems using platinum electrode as electron souce. The use of electrode systems for P450 catatlysed reactions can be utilized for deveoplement of biosensors and bioelectronics.
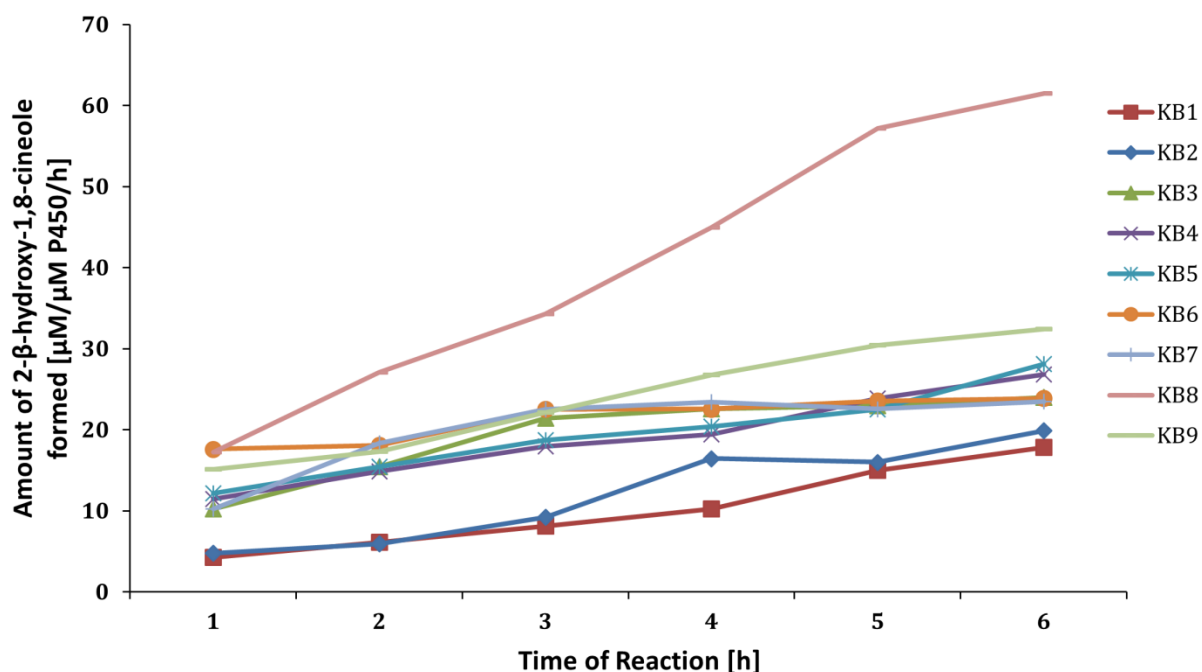
Figure 4.6.6. 2-β-hydroxy-1,8-cineole formed by P450cin variants in electrobiocatalytic reaction set up.
2-β-hydroxy-1,8-cineole formed by P450cin variants in electrobiocatalytic reaction employing Platinum electrodes as electron source and Cobalt(III)sepulchrate as mediator using 1 μM P450, 6 mM 1,8-cineole, 5 mM Cobalt(III)sepulchrate. Product formation was determined against -700 mV vs Ag/AgCl in a reaction time of 6 h. WT (CinA-linker10aa-CinC); KB1(Q385H); KB2 (V386S); KB3 (Q385H/V386S); KB4 (Q385H/V386S/T77C); KB5 (Q385H/V386S/L88N); KB6 (Q385H/V386S/T77S/L88C); KB7 (Q385H/V386S/T77C/L88Y); KB8 (Q385H/V386S/T77N/L88R); KB9 (Q385H/V386S/P79G/R80D/Y81S).

Maximum 2-β-hydroxy-1,8-cineole formation was observed for mutant KB8 (61.5 ± 1.09 μM μM^{-1} P450) after a reaction of 6 h which was approximately 3 times higher than the WT variant (23.51± 2.5 μM μM^{-1} P450) (Figure 4.6.6). The rate of product formation for variant KB8 (0.16 nmol (nmol) P450^{-1} min^{-1} cm^{-2}) was also 3 folds higher than the WT variant (0.038 nmol (nmol) P450^{-1} min^{-1} cm^{-2}). Control reaction without any Cobalt(III)sepulchrate and without any application of potential and using crude cell lysate of empty vector (pALXtreme-1a) did not shown 2-β-hydroxy-1,8-cineole formation.
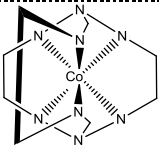
4.6.8. Discussion

Electrobiocatalytic mediated conversion of substrates is a well-known approach for substitution of NAD(P)H in P450 group of enzymes (Fleming et al., 2006; Hollmann et al., 2006; Holtmann and Schrader, 2007; Shumyantseva et al., 2007; Shumyantseva et al., 2005). In most of the studies previously reported a mediator was used in combination with electrodes in bioelectrochemical reaction set up. Mediators are artificial electron transferring agents which shuttle electrons from the surface of an electrode to the redox center of enzymes. A suitable mediator should fulfil several requirements (Holtmann and Schrader, 2007):

- It should be stable under required working conditions and should not participate in the side reactions during electron transfer.
- It should show electrochemically reversible reaction kinetics and should be stable under oxidized and reduced forms.
- It must be soluble under reaction conditions.
- The redox potential of the mediator should be more negative than the redox potential of the enzyme's active site for reductive biocatalysis.

The mediator Cobalt(III)sepulchrate has previously been described to be used in P450 BM 3 and P450cam catalyzed electrobiocatalytic reactions. (Faulkner et al., 1995; Nazor et al., 2008; Nazor and Schwaneberg, 2006; Shumyantseva et al., 2005). Cobalt(III)sepulchrate was suitable for use in cytochrome P450cin catalyzed reactions since the redox potential of Cobalt(III)sepulchrate is more negative than P450cin and CinC (cindoxin) (Table 4.6.1). Also, it is soluble under the used reaction conditions.

Table 4.6.1. Standard potential of Co (III) sep , P450cin and Cindoxin vs Ag/AgCl.

Mediator/prosthetic group/protein	Standard potential vs Ag/AgCl [mV]	Structure	Reference
P450cin	-402	-	Kimmich et al., 2007
Cindoxin	-430	-	Kimmich et al., 2007
Cobalt (III) sepulchrate	-615		Anderson et al., 1997

With a limited screening capacity using a potentiostat the screening of the whole P450cin mutant library using the above said electrochemical set up was not feasible. So screening of P450cin variants which displayed an improvement in activity for Zinc dust and Cobalt(III)sepulchrate was performed of which mutant KB8 showed the highest product formation (61 µM µM^{-1} P450) in a reaction time of 6 h with Platinum and Cobalt(III)sepulchrate (Figure 4.6.2).

Cekic et al., reported a conversion rate of 6.50 nmol (nmol) P450^{-1} min^{-1} cm^{-2} using CinA, CinC (ratio 1:3) and Cobalt(III)sepulchrate (5 mM) (Çekiç et al., 2010). Using mutant KB8 (Q385H/V386S/T77N/L88R) conversion rate of 0.16 nmol (nmol) P450^{-1} min^{-1} cm^{-2} which was approximately 3 times higher than the WT variant 0.038 nmol (nmol) P450^{-1} min^{-1} cm^{-2} was obtained. The turnover rates using fusion protein reached only about 2% of the activity as reported by Cekic et al., for reconstituted cytochrome P450cin system consisting of CinA, CinC and *E. coli* Fpr (ratio 1:8:2).

One possible reason for low activity of the fusion protein employing the electrochemical set up could be the stability of the fusion protein or the fused CinA and CinC might not be optimally orientated with respect to each for use in electrochemical reaction set up. Experimental evidence support that CinA and CinC lose their activity after 5 and 8 h respectively in the fusion protein. Cekic et al reported 80% and 20% of initial activity for CinA after 4 h and 24 h respectively under identical electrochemical reaction condition used in this study (Çekiç et al., 2010). Cobalt(III)sepulchrate is highly stable in both oxidized and reduced forms. The peak currents were 70% after 500 CVs compared to the initial cycle (Cekic 2012, PhD thesis, RWTH Aachen University). Subsequently, external addition of

crude cell lysate of CinA and CinC in the reaction mixture did not significantly increase the activity of the fusion protein. This suggests that the stability and orientation of protein are not the only factors which can contribute to the productivity of an enzyme system in an electrobiocatalytic reaction. There are numerous other factors to be considered like availability of oxygen, potential applied, surface area of electrode, time of reaction which significantly contributes to the formation of product in an electrobiocatalytic reaction system.

The comparison of 2-β-hydroxy-1,8-cineole formed using P450cin variants employing NADPH, Zinc dust and Cobalt(III)sepulchrate and electrobiocatalytic reaction set up is as shown in figure 4.6.7. Activity of the P450cin variants using electrochemical approach is still far less compared to the activity obtained using NADPH as well as Zinc dust and Cobalt(III)sepulchrate.
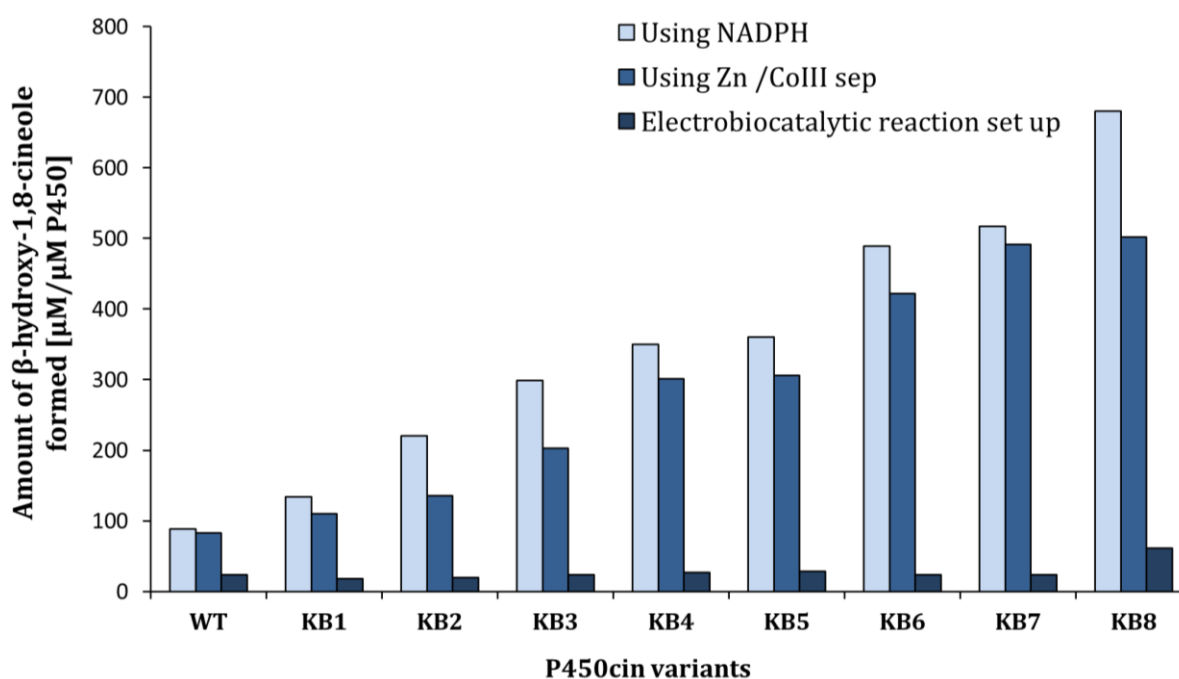
Figure 4.6.7. Comparison of activity of P450cin variants using NADPH, Zinc dust and Cobalt(III)sepulchrate and electrobiocatalytic reaction set up. WT (CinA-linker10aa-CinC); KB1(Q385H); KB2 (V386S); KB3 (Q385H/V386S); KB4 (Q385H/V386S/T77C); KB5 (Q385H/V386S/L88N); KB6 (Q385H/V386S/T77S/L88C); KB7 (Q385H/V386S/T77C/L88Y); KB8 (Q385H/V386S/T77N/L88R).

Recent report by Tosstorff et al., describe the effect of molecular oxygen on oxidation-reduction of cobalt-sepulchrate. A difficulty occurring during the use of cobalt(III)sepulchrate in electrobiocatalytic reactions is the oxidative uncoupling of the mediator oxidation as molecular oxygen can re-oxidize the mediator (Tosstorff et al., 2014). Another challenge while developing an enzyme for electrobiocatalytic application is a high throughput screening system. Screening using a combination of electromicrotiter plates (eMTPs) (Ley et al., 2013) or colorimetric or fluorometric assays based on change in oxidation state of mediator can provide the necessary throughput for screening of mutants for electrobiocatalytic properties. Additionally improved set-up of the electrochemical cell, e.g. by use of packed bed flow cell can increase the productivity of the cytochrome P450cin system.

4.6.9. Conclusion

In summary, preliminary electrochemical investigation of selected P450cin variants (KB1-KB8), identified KB8 which showed the highest product formation (61 μM μM^{-1} P450) in a reaction time of 6 h with Platinum and Cobalt(III)sepulchrate. Using mutant KB8 (Q385H/V386S/T77N/L88R) conversion rate of 0.16 nmol (nmol) P450^{-1} min^{-1} cm^{-2} which is approximately 3 times higher than the WT variant 0.038 nmol (nmol) P450^{-1} min^{-1} cm^{-2} was obtained. The activity and productivity of the P450cin variants is still lower by a magnitude of about 6-7 folds lower than for NADPH. Nonetheless, this study shows that it is possible to redign the P450 monooxygenases for electrobiocatalytic applications using directed evolution strategy.

5. SUMMARY AND CONCLUSION

Over the past decade, protein engineering has been used for the improvement of various enzyme classes like proteases, lacasses, glucose oxidase, cytochrome P450s for industrial applications. Protein engineering studies also provide useful structural and functional insights for development of newer and innovative biocatalysts.

The work described in this thesis tackles two major hurdles in the application of cytochrome P450cin for industrial biocatalysis using protein engineering approach. The two obstacles of using cytochrome P450cin are- first the complex multicomponent nature of P450cin and second the requirement of NADPH, an expensive cofactor for hydroxylation of its substrate 1,8-cineole.

The first part of the thesis describes a novel, robust, technically simple and enzyme free method (P-LinK-Protein fusion with variable linker insertion) for generating fusion proteins with variable linker length. P-Link was successfully applied for fusion of CinC domain to the C terminal end of CinA of P450cin system through sixteen different linker lengths in a single experiment employing three PCRs. The resulting CinA-CinC linker library was screened using a TLC based screening system for detection of 2-β-hydroxy-1,8-cineole to identify the most active fusion protein construct. Linker lengths between 8 to 12 amino acids were determined to be the most optimal for CinA and CinC fusion protein with a maximum observed for fusion protein with a linker length of 10 amino acids. Fusion proteins with a linker length of 4 amino acids or less were determined to be inactive. Although the fusion protein generated in this study shows lower activity (176 μM μM^{-1} P450) compared to equimolar mixture of its individual component (315 μM μM^{-1} P450), the use of fusion proteins simplifies the multicomponent system on genetic level and enables its use in protein engineering campaigns. P-Link can also be employed to vary the amino acid sequence within linkers (e.g. two positions with NNK codons) in the primer sequence can be incorporated to explore and tune flexibility of the linker sequence and thereby increase activity of the fusion protein. It is further likely that P-LinK can be expanded iteratively to at least three component systems in which a third electron transfer partner of monooxygenase systems can be fused.

The P-LinK method is the first reported method which can be used to fuse two or more protein domains and also vary the length of the linker between the fused domains in a

143

single experiment. P-Link is not only limited to P450 enzymes and can be of general interest to generate fusion enzymes to perform, for instance, cascade reactions. The use of P-Link method simplifies multicomponent enzyme system on genetic level and enables its use in protein engineering.

Subsequently, characterization of the fusion protein with 10 amino acid linker length was performed using computational and experimental approach. It was performed in order to elucidate the electron transfer pathway and the role of linker length in the activity of the fusion protein. Individual mutants FMN (-) with mutation G594D in the FMN domain and heme (-) with a mutation A91F in the heme domain were determined to be inactive and did not show formation of 2-β-hydroxy-1,8-cineole. However, for an equimolar mixture of individual inactive domains (Heme (-) + FMN (-)) formation of 2-β-hydroxy-1,8-cineole was detected. This suggests that the electron transfer in the CinA-CinC fusion protein occurs predominantly in an intermolecular manner from heme domain of one molecule to the FMN domain of the other molecule in the dimeric P450cin fusion protein. This is the first report in which the fusion protein of CinA and CinC (a cytochrome P450 from plant origin) has been reported to be dimeric in structure and the electron transfer pathway described to be from heme domain of one molecule to the FMN domain of the other molecule in the dimeric protein.

In order to engineer the P450cin fusion protein (CinA-linker 10 aa-CinC) for alternative electron transfer system (Zinc dust and Cobalt(III)sepulchrate) two rounds of random mutagenesis libraries were generated with the SeSaM method. Screening of SeSaM libraries using a validated TLC based screening system determined four positions Q385, V386, T77 and L88 to be crucial for improvement in activity towards 1,8-cineole using Zinc dust and Cobalt(III)sepulchrate as an alternative electron transfer system. A final variant KB8 (Q385H/V386S/T77N/L88R) was identified which exhibited a catalytic efficiency (k_{eff} = 27.77 μM min^{-1} mM^{-1}) (k_{cat}/K$_m$) 3.8 times higher as the starting variant (CinA-linker10aa-CinC) (k_{eff} = 7.4 μM min^{-1} mM^{-1}). Variant KB8 showed a 3.7 fold increase in k_{cat} value (2.5 μM min^{-1}) in comparison to the WT (0.67 μM min^{-1}) with no significant change in the k_m value (90 μM). Additionally the total turnover numbers were determined under non optimized conditions for Zinc and Cobalt(III)sepulchrate to be 1170 and 480 for mutant KB8 using 5 mM and 1 mM 1,8-cineole concentration respectively.

144

A semi-rational approach was used to rationally select amino acid positions 79, 80, and 81 in P450cin fusion protein after docking of Cobalt(III)sepulchrate in the active site of the enzyme. Residues 79, 80, and 81 were selected as they were in close proximity to the amino acid positions identified by random mutagenesis (77 and 88) and were in direct contact with Cobalt(III)sepulchrate. In this way, residues 79, 80, and 81 could complete a putative 'path' between Cobalt(III)sepulchrate and heme moiety in the active site of the enzyme. Screening of simultaneous site saturation mutagenesis library identified a variant KB9 (Q385H/V386S/P79G/R80D/Y81S) with about 1.5 times increased hydroxylation activity compared to variant KB3 (Q385H/V386S). This is a first report describing positions which were identified from random mutant libraries (amino acid residues Q385, V386, T77, L88) and rationally designed mutant libraries (amino acid residues 79, 80, and 81) using an alternative cofactor system. The amino acid residues and the substitution pattern identified in this work can be further extrapolated to other P450s for development of improved P450 catalysts.

As a final step, preliminary electrochemical investigations of selected P450cin variants (KB1- KB8) were performed. Mutant KB8 showed the highest product formation (61 µM µM^{-1} P450) in a reaction time of 6 h with Platinum and Cobalt(III)sepulchrate. Product formation rate of 0.16 nmol (nmol) P450^{-1} min^{-1} cm^{-2} which is approximately 3 times higher than the WT variant 0.038 nmol (nmol) P450^{-1} min^{-1} cm^{-2} was obtained in a total reaction time of 6 h.

Results obtained in this work show that efficient P450cin fusion constructs can be designed and optimized by directed evolution and semi-rational design. The developed concepts of screening system, the P-Link method for construction of fusion proteins and identified amino acid residues build a foundation for potential application of P450s in biosensors, biofuels and biocatalysis.

6. APPENDIX

6.1. List of Figures

6.2. List of Tables

6.3. List of Schemes

6.4. List of Primers

Chapter 2

Table 1. List of nucleotides used for construction of SeSaM libraries.

Primer Name	Sequence	Purpose
P1	cgcctgtcaccgactcactatagggaattgtgagcgga	Template generation_SeSaM step 1
P2	gtgtgatggcgtgaggcagccgggctttgttagcagccggatctcag	Template generation_SeSaM step 1
P3	cacactaccgcactccgtcgcgactcactatagggaattgtgagcgga	Template generation_SeSaM step 3
P4	gcggacagtgcgggctttgttagcagccggatctcag	Template generation_SeSaM step 3
P5	**ctaacaaagcccg**aaaggaagctgagttg	Cloning of SeSaM library_Fwd_gene amplification
P6	**cgggctttgttag**cagccggatctcag	Cloning of SeSaM library_Rev_gene amplification
P7	**gactcactatag**gggaattgtgagcgg	Cloning of SeSaM library_Fwd_vector amplification
P8	**ctatagtgagtc**gtattaatttcgaacatgtgagc	Cloning of SeSaM library_Rev_vector amplification

* Bold and underlined letters indicate phosphorothioate bases.

Chapter 4

Section 4.2

Table 2A. List of nucleotides used for construction of CinA-CinC fusion protein.

Primer Name	Sequence	Purpose
FA1	**gactcactatag**gggaattgtgagcgg	*cinA* amplification
RA1	**ttcgctcagacg**tttgcctttcgg	*cinA* amplification
FC1	**cgtctgagcgaa**atgggtaatgccctgattttatatggc	*cinC* amplification
RC1	**cgggctttgttag**cagccggatctcag	*cinC* amplification
FP1	**ctaacaaagcccg**aaaggaagctgagttg	Vector amplification
RP1	**ctatagtgagtc**gtattaatttcgaacatgtgagc	Vector amplification

* Bold and underlined letters indicate phosphorothioate bases.

Table 2B. List of nucleotides used for insertion of linkers (length 1-16) in the CinA-CinC fusion protein.

Primer Name	Purpose	Sequence
RL1	R.P. for linkers 1-5	**ttcgctcagacg**TTTGCCTTTCGGAAAAATAATC
FL1	F.P. for Linker 1	**cgtctgagcgaa**CCTATGGGTAATGCCCTGATTTTATATGGCAC
FL2	F.P. for Linker 2	**cgtctgagcgaa**CCTTCTATGGGTAATGCCCTGATTTTATATGGCAC
FL3	F.P. for Linker 3	**cgtctgagcgaa**CCTTCTCCAATGGGTAATGCCCTGATTTTATATGGCAC
FL4	F.P. for Linker 4	**cgtctgagcgaa**CCTTCTCCATCTATGGGTAATGCCCTGATTTTATATGGC
FL5	F.P. for Linker 5	**cgtctgagcgaa**CCTTCTCCATCTACTATGGGTAATGCCCTGATTTTATATGGC
RL2	R.P. for linkers 6-11	**agtagatggaga**AGGTTCGCTCAGACGTTTGCCTTTC
FL6	F.P. for Linker 6	**tctccatctact**GACATGGGTAATGCCCTGATTTTATATGGCAC
FL7	F.P. for Linker 7	**tctccatctact**GACCAAATGGGTAATGCCCTGATTTTATATGGCAC
FL8	F.P. for Linker 8	**tctccatctact**GACCAATCCATGGGTAATGCCCTGATTTTATATGGCAC
FL9	F.P. for Linker 9	**tctccatctact**GACCAATCCCCTATGGGTAATGCCCTGATTTTATATGGCAC
FL10	F.P. for Linker 10	**tctccatctact**GACCAATCCCCTTCTATGGGTAATGCCCTGATTTTATATG
FL11	F.P. for Linker 11	**tctccatctact**GACCAATCCCCTTCTACTATGGGTAATGCCCTGATTTTATATG
RL3	R.P. for linkers 12-16	**agtagaaggggga**TTGGTCAGTAGATGGAGAAGGTTCGCTCAGACGTTTGCC
FL12	F.P. for Linker 12	**tccccttctact**GGAATGGGTAATGCCCTGATTTTATATGGCAC
FL13	F.P. for Linker 13	**tccccttctact**GGAGACATGGGTAATGCCCTGATTTTATATGGCAC
FL14	F.P. for Linker 14	**tccccttctact**GGAGACGCTATGGGTAATGCCCTGATTTTATATGGCAC
FL15	F.P. for Linker 15	**tccccttctact**GGAGACGCTGTTATGGGTAATGCCCTGATTTTATATG
FL16	F.P. for Linker 16	**tccccttctact**GGAGACGCTGTTGCTATGGGTAATGCCCTGATTTTATATG

* Bold and underlined letters indicate phosphorothioate bases.

Section 4.3

Table 3. List of nucleotides used for construction of SDM mutants A91F (Heme (-)) and G594D (FMN (-))

Primer Name	Primer sequence	Purpose
SDM_A91F_FP	gaactgatgatgtttggccaggatgatccg	Forward primer for SDM at position A91
SDM_A91F_RP	cggatcatcctggccaaacatcatcagttc	Reverse primer for SDM at position A91
SDM_G594D_FP	gtgtttggtctggatgatagctattacactacg	Forward primer for SDM at position G594
SDM_G594D_RP	cgtagtgtaatagctatcatccagaccaaacac	Reverse primer for SDM at position G594

Section 4.4

Table 4. List of nucleotides used for construction of SSM libraries at position Q385, V386, T77 and L88

Primer Name	Sequence	Purpose
SSM_Q385V386_FP	gctgatgggcnnknnkgcgggcatgctgc	Forward primer for SSM at position 385 and 386
SSM_Q385V386_RP	gcagcatgcccgcmnnmnngcccatcagc	Forward primer for SSM at position 385 and 386
SSM_77_FP	gcaacaaaggcgtgnnktttccgcgttatgaaaccg	Forward primer for SSM at position 77
SSM_77_RP	cggtttcataacgcggaaamnncacgcctttgttgc	Reverse primer for SSM at position 77
SSM_88_FP	ccggcgaatttgaannkatgatggcgggccag	Forward primer for SSM at position 88
SSM_88_RP	ctggcccgccatcatmnnttcaaattcgccgg	Reverse primer for SSM at position 88

Section 4.5

Table 5. List of nucleotides used for construction of simultaneous SSM library at position P79, R80 and Y81.

Primer Name	Primer sequence	Purpose
SSM_798081_FP	ggcgtgacctttnnknnknnkgaaaccggcgaatttg	Forward primer for SSM at position 79,80,81
SSM_798081_RP	caaattcgccggtttcmnnmnnmnnaaaggtcacgcc	Reverse primer for SSM at position 79,80,81

Protein and Nucleotide sequences

CinA WT

Start codon

atg ggc acc gcg acc gtg gcg agc acc agc ctg ttt acc acc gcg gat cat tat
M G   T   A   T   V   A   S   T   S   L   F   T   T   A   D   H   Y

cat acc ccg ctg ggc ccg gat ggc acc ccg cat gcg ttt ttt gaa gcg ctg cgt
H   T   P   L   G   P   D   G   T   P   H   A   F   F   E   A   L   R

gat gaa gcg gaa acc acc ccg att ggt tgg agc gaa gcg tat ggc ggc cat tgg
D   E   A   E   T   T   P   I   G   W   S   E   A   Y   G   G   H   W

gtg gtg gcg ggc tat aaa gaa att cag gcg gtg att cag aac acc aaa gcg ttt
V   V   A   G   Y   K   E   I   Q   A   V   I   Q   N   T   K   A   F

agc aac aaa ggc gtg acc ttt ccg cgt tat gaa acc ggc gaa ttt gaa ctg atg
S   N   K   G   V   T   F   P   R   Y   E   T   G   E   F   E   L   M

atg gcg ggc cag gat gat ccg gtg cat aaa aaa tat cgt cag ctg gtt gcg aaa
M   A   G   Q   D   D   P   V   H   K   K   Y   R   Q   L   V   A   K

ccg ttt agc ccg gaa gcg acc gac ctg ttt acc gaa cag ctg cgt cag agc acc
P   F   S   P   E   A   T   D   L   F   T   E   Q   L   R   Q   S   T

aac gat ctg att gat gcg cgt att gaa ctg ggc gaa ggt gat gcg gcg acc tgg
N   D   L   I   D   A   R   I   E   L   G   E   G   D   A   A   T   W

ctg gcc aac gaa att ccg gcg cgt ctg acc gcg att ctg ctg ggt ctg ccg ccg
L   A   N   E   I   P   A   R   L   T   A   I   L   L   G   L   P   P

gaa gat ggc gat acc tat cgt cgt tgg gtg tgg gcg att acc cat gtg gaa aat
E   D   G   D   T   Y   R   R   W   V   W   A   I   T   H   V   E   N

ccg gaa gaa ggc gcg gaa att ttt gcg gaa ctg gtg gcg cac gcg cgt acc ctg
P   E   E   G   A   E   I   F   A   E   L   V   A   H   A   R   T   L

att gcg gaa cgt cgt acc aat ccg ggc aac gat att atg agc cgt gtg atc atg
I   A   E   R   R   T   N   P   G   N   D   I   M   S   R   V   I   M

agc aaa att gat ggc gaa agc ctg agc gaa gat gat ctg att ggc ttt ttt acc
S   K   I   D   G   E   S   L   S   E   D   D   L   I   G   F   F   T

att ctg ctg ctg ggc ggc att gat aac acc gcg cgt ttt ctg agc agc gtg ttt
I   L   L   L   G   G   I   D   N   T   A   R   F   L   S   S   V   F

tgg cgt ctg gcc tgg gat att gaa ctg cgt cgt cgt ctg att gcg cat ccg gaa
W   R   L   A   W   D   I   E   L   R   R   R   L   I   A   H   P   E

ctg att ccg aac gcg gtg gat gaa ctg ctg cgt ttt tat ggt ccg gca atg gtg
L   I   P   N   A   V   D   E   L   L   R   F   Y   G   P   A   M   V

ggt cgt ctg gtg acc cag gaa gtg acc gtt ggc gat att acc atg aaa ccg ggc
G   R   L   V   T   Q   E   V   T   V   G   D   I   T   M   K   P   G

```
cag acc gcc atg ctg tgg ttt ccg att gcg agc cgt gat cgt agc gcg ttt gat
 Q   T   A   M   L   W   F   P   I   A   S   R   D   R   S   A   F   D
agc ccg gat aac att gtg att gaa cgt acc ccg aac cgt cat ctg agc ctg ggt
 S   P   D   N   I   V   I   E   R   T   P   N   R   H   L   S   L   G

cat ggc att cat cgt tgc ctg ggc gcg cat ctg att cgt gtg gaa gcg cgt gtg
 H   G   I   H   R   C   L   G   A   H   L   I   R   V   E   A   R   V

gcg att acc gaa ttt ctg aaa cgc att ccg gaa ttt agc ctg gac ccg aat aaa
 A   I   T   E   F   L   K   R   I   P   E   F   S   L   D   P   N   K

gaa tgc gaa tgg ctg atg ggc cag gtg gcg ggc atg ctg cat gtg ccg att att
 E   C   E   W   L   M   G   Q   V   A   G   M   L   H   V   P   I   I

ttt ccg aaa ggc aaa cgt ctg agc gaa taa taa
 F   P   K   G   K   R   L   S   E   -   -
                                    Stop codon
```

CinC WT

Start codon

```
atg ggt aat gcc ctg att tta tat ggc acg gag act ggt aat gcg gag gcg tgt
 M   G   N   A   L   I   L   Y   G   T   E   T   G   N   A   E   A   C

gcg acc aca att tcg cag gtt ctg gcg gat acc gtt gat acg aag gtg cat gat
 A   T   T   I   S   Q   V   L   A   D   T   V   D   T   K   V   H   D

ctg gcc gat atg acc ccc cgc gca atg ctg gat agt ggt gcg gat ctg atc gtg
 L   A   D   M   T   P   R   A   M   L   D   S   G   A   D   L   I   V

ttt gcc acg gca acc tat ggc gaa ggc gaa ttt gcg ggt ggt ggc gca gcg ttc
 F   A   T   A   T   Y   G   E   G   E   F   A   G   G   G   A   A   F

ttt gag act ctg cgt gaa acc aaa ccc gat ctg agc ggt ctg cgt ttc gct gtg
 F   E   T   L   R   E   T   K   P   D   L   S   G   L   R   F   A   V

ttt ggt ctg ggc gat agc tat tac act acg ttt aac cag gcc ggt gca act gcc
 F   G   L   G   D   S   Y   Y   T   T   F   N   Q   A   G   A   T   A

gca acc att tta gcc agc ctg ggt ggt acc cag gtg ggt gat acg gcg cgc cac
 A   T   I   L   A   S   L   G   G   T   Q   V   G   D   T   A   R   H

gat act tcg tcg ggc gac gat cca gaa gaa acc gcc gaa gaa tgg gcc cgc gag
 D   T   S   S   G   D   D   P   E   E   T   A   E   E   W   A   R   E

atc ttg act gca ttg gca act ccg gca gtt agt taa taa
 I   L   T   A   L   A   T   P   A   V   S   -   -
                                             Stop codon
```

CinA-CinC fusion (linker length 0)

Start codon

```
atg ggc acc gcg acc gtg gcg agc acc agc ctg ttt acc acc gcg gat cat tat
 M   G   T   A   T   V   A   S   T   S   L   F   T   T   A   D   H   Y

cat acc ccg ctg ggc ccg gat ggc acc ccg cat gcg ttt ttt gaa gcg ctg cgt
 H   T   P   L   G   P   D   G   T   P   H   A   F   F   E   A   L   R

gat gaa gcg gaa acc acc ccg att ggt tgg agc gaa gcg tat ggc ggc cat tgg
 D   E   A   E   T   T   P   I   G   W   S   E   A   Y   G   G   H   W

gtg gtg gcg ggc tat aaa gaa att cag gcg gtg att cag aac acc aaa gcg ttt
 V   V   A   G   Y   K   E   I   Q   A   V   I   Q   N   T   K   A   F

agc aac aaa ggc gtg acc ttt ccg cgt tat gaa acc ggc gaa ttt gaa ctg atg
 S   N   K   G   V   T   F   P   R   Y   E   T   G   E   F   E   L   M

atg gcg ggc cag gat gat ccg gtg cat aaa aaa tat cgt cag ctg gtt gcg aaa
 M   A   G   Q   D   D   P   V   H   K   K   Y   R   Q   L   V   A   K

ccg ttt agc ccg gaa gcg acc gac ctg ttt acc gaa cag ctg cgt cag agc acc
 P   F   S   P   E   A   T   D   L   F   T   E   Q   L   R   Q   S   T

aac gat ctg att gat gcg cgt att gaa ctg ggc gaa ggt gat gcg gcg acc tgg
 N   D   L   I   D   A   R   I   E   L   G   E   G   D   A   A   T   W

ctg gcc aac gaa att ccg gcg cgt ctg acc gcg att ctg ctg ggt ctg ccg ccg
 L   A   N   E   I   P   A   R   L   T   A   I   L   L   G   L   P   P

gaa gat ggc gat acc tat cgt cgt tgg gtg tgg gcg att acc cat gtg gaa aat
 E   D   G   D   T   Y   R   R   W   V   W   A   I   T   H   V   E   N

ccg gaa gaa ggc gcg gaa att ttt gcg gaa ctg gtg gcg cac gcg cgt acc ctg
 P   E   E   G   A   E   I   F   A   E   L   V   A   H   A   R   T   L

att gcg gaa cgt cgt acc aat ccg ggc aac gat att atg agc cgt gtg atc atg
 I   A   E   R   R   T   N   P   G   N   D   I   M   S   R   V   I   M

agc aaa att gat ggc gaa agc ctg agc gaa gat gat ctg att ggc ttt ttt acc
 S   K   I   D   G   E   S   L   S   E   D   D   L   I   G   F   F   T

att ctg ctg ctg ggc ggc att gat aac acc gcg cgt ttt ctg agc agc gtg ttt
 I   L   L   L   G   G   I   D   N   T   A   R   F   L   S   S   V   F

tgg cgt ctg gcc tgg gat att gaa ctg cgt cgt cgt ctg att gcg cat ccg gaa
 W   R   L   A   W   D   I   E   L   R   R   R   L   I   A   H   P   E

ctg att ccg aac gcg gtg gat gaa ctg ctg cgt ttt tat ggt ccg gca atg gtg
 L   I   P   N   A   V   D   E   L   L   R   F   Y   G   P   A   M   V

ggt cgt ctg gtg acc cag gaa gtg acc gtt ggc gat att acc atg aaa ccg ggc
 G   R   L   V   T   Q   E   V   T   V   G   D   I   T   M   K   P   G

cag acc gcc atg ctg tgg ttt ccg att gcg agc cgt gat cgt agc gcg ttt gat
```

```
  Q   T   A   M   L   W   F   P   I   A   S   R   D   R   S   A   F   D
agc ccg gat aac att gtg att gaa cgt acc ccg aac cgt cat ctg agc ctg ggt
  S   P   D   N   I   V   I   E   R   T   P   N   R   H   L   S   L   G

cat ggc att cat cgt tgc ctg ggc gcg cat ctg att cgt gtg gaa gcg cgt gtg
  H   G   I   H   R   C   L   G   A   H   L   I   R   V   E   A   R   V

gcg att acc gaa ttt ctg aaa cgc att ccg gaa ttt agc ctg gac ccg aat aaa
  A   I   T   E   F   L   K   R   I   P   E   F   S   L   D   P   N   K

gaa tgc gaa tgg ctg atg ggc cag gtg gcg ggc atg ctg cat gtg ccg att att
  E   C   E   W   L   M   G   Q   V   A   G   M   L   H   V   P   I   I

ttt ccg aaa ggc aaa cgt ctg agc gaa atg ggt aat gcc ctg att tta tat ggc
  F   P   K   G   K   R   L   S   E   M   G   N   A   L   I   L   Y   G
                                          CinC gene starts
acg gag act ggt aat gcg gag gcg tgt gcg acc aca att tcg cag gtt ctg gcg
  T   E   T   G   N   A   E   A   C   A   T   T   I   S   Q   V   L   A

gat acc gtt gat acg aag gtg cat gat ctg gcc gat atg acc ccc cgc gca atg
  D   T   V   D   T   K   V   H   D   L   A   D   M   T   P   R   A   M

ctg gat agt ggt gcg gat ctg atc gtg ttt gcc acg gca acc tat ggc gaa ggc
  L   D   S   G   A   D   L   I   V   F   A   T   A   T   Y   G   E   G

gaa ttt gcg ggt ggt ggc gca gcg ttc ttt gag act ctg cgt gaa acc aaa ccc
  E   F   A   G   G   G   A   A   F   F   E   T   L   R   E   T   K   P

gat ctg agc ggt ctg cgt ttc gct gtg ttt ggt ctg ggc gat agc tat tac act
  D   L   S   G   L   R   F   A   V   F   G   L   G   D   S   Y   Y   T

acg ttt aac cag gcc ggt gca act gcc gca acc att tta gcc agc ctg ggt ggt
  T   F   N   Q   A   G   A   T   A   A   T   I   L   A   S   L   G   G

acc cag gtg ggt gat acg gcg cgc cac gat act tcg tcg ggc gac gat cca gaa
  T   Q   V   G   D   T   A   R   H   D   T   S   S   G   D   D   P   E

gaa acc gcc gaa gaa tgg gcc cgc gag atc ttg act gca ttg gca act ccg gca
  E   T   A   E   E   W   A   R   E   I   L   T   A   L   A   T   P   A

gtt agt taa taa
  V   S   -   -
        Stop codon
```

157

CinA-CinC fusion (linker length 10 amino acids)

Start codon

```
atg ggc acc gcg acc gtg gcg agc acc agc ctg ttt acc acc gcg gat cat tat
 M   G   T   A   T   V   A   S   T   S   L   F   T   T   A   D   H   Y

cat acc ccg ctg ggc ccg gat ggc acc ccg cat gcg ttt ttt gaa gcg ctg cgt
 H   T   P   L   G   P   D   G   T   P   H   A   F   F   E   A   L   R

gat gaa gcg gaa acc acc ccg att ggt tgg agc gaa gcg tat ggc ggc cat tgg
 D   E   A   E   T   T   P   I   G   W   S   E   A   Y   G   G   H   W

gtg gtg gcg ggc tat aaa gaa att cag gcg gtg att cag aac acc aaa gcg ttt
 V   V   A   G   Y   K   E   I   Q   A   V   I   Q   N   T   K   A   F

agc aac aaa ggc gtg acc ttt ccg cgt tat gaa acc ggc gaa ttt gaa ctg atg
 S   N   K   G   V   T   F   P   R   Y   E   T   G   E   F   E   L   M

atg gcg ggc cag gat gat ccg gtg cat aaa aaa tat cgt cag ctg gtt gcg aaa
 M   A   G   Q   D   D   P   V   H   K   K   Y   R   Q   L   V   A   K

ccg ttt agc ccg gaa gcg acc gac ctg ttt acc gaa cag ctg cgt cag agc acc
 P   F   S   P   E   A   T   D   L   F   T   E   Q   L   R   Q   S   T

aac gat ctg att gat gcg cgt att gaa ctg ggc gaa ggt gat gcg gcg acc tgg
 N   D   L   I   D   A   R   I   E   L   G   E   G   D   A   A   T   W

ctg gcc aac gaa att ccg gcg cgt ctg acc gcg att ctg ctg ggt ctg ccg ccg
 L   A   N   E   I   P   A   R   L   T   A   I   L   L   G   L   P   P

gaa gat ggc gat acc tat cgt cgt tgg gtg tgg gcg att acc cat gtg gaa aat
 E   D   G   D   T   Y   R   R   W   V   W   A   I   T   H   V   E   N

ccg gaa gaa ggc gcg gaa att ttt gcg gaa ctg gtg gcg cac gcg cgt acc ctg
 P   E   E   G   A   E   I   F   A   E   L   V   A   H   A   R   T   L

att gcg gaa cgt cgt acc aat ccg ggc aac gat att atg agc cgt gtg atc atg
 I   A   E   R   R   T   N   P   G   N   D   I   M   S   R   V   I   M

agc aaa att gat ggc gaa agc ctg agc gaa gat gat ctg att ggc ttt ttt acc
 S   K   I   D   G   E   S   L   S   E   D   D   L   I   G   F   F   T

att ctg ctg ctg ggc ggc att gat aac acc gcg cgt ttt ctg agc agc gtg ttt
 I   L   L   L   G   G   I   D   N   T   A   R   F   L   S   S   V   F

tgg cgt ctg gcc tgg gat att gaa ctg cgt cgt cgt ctg att gcg cat ccg gaa
 W   R   L   A   W   D   I   E   L   R   R   R   L   I   A   H   P   E

ctg att ccg aac gcg gtg gat gaa ctg ctg cgt ttt tat ggt ccg gca atg gtg
 L   I   P   N   A   V   D   E   L   L   R   F   Y   G   P   A   M   V

ggt cgt ctg gtg acc cag gaa gtg acc gtt ggc gat att acc atg aaa ccg ggc
 G   R   L   V   T   Q   E   V   T   V   G   D   I   T   M   K   P   G

cag acc gcc atg ctg tgg ttt ccg att gcg agc cgt gat cgt agc gcg ttt gat
```

```
 Q   T   A   M   L   W   F   P   I   A   S   R   D   R   S   A   F   D
agc ccg gat aac att gtg att gaa cgt acc ccg aac cgt cat ctg agc ctg ggt
 S   P   D   N   I   V   I   E   R   T   P   N   R   H   L   S   L   G

cat ggc att cat cgt tgc ctg ggc gcg cat ctg att cgt gtg gaa gcg cgt gtg
 H   G   I   H   R   C   L   G   A   H   L   I   R   V   E   A   R   V

gcg att acc gaa ttt ctg aaa cgc att ccg gaa ttt agc ctg gac ccg aat aaa
 A   I   T   E   F   L   K   R   I   P   E   F   S   L   D   P   N   K

gaa tgc gaa tgg ctg atg ggc cag gtg gcg ggc atg ctg cat gtg ccg att att
 E   C   E   W   L   M   G   Q   V   A   G   M   L   H   V   P   I   I

ttt ccg aaa ggc aaa cgt ctg agc gaa cct tct cca tct act gac caa tcc cct
 F   P   K   G   K   R   L   S   E   P   S   P   S   T   D   Q   S   P
                                        Linker sequence
tct atg ggt aat gcc ctg att tta tat ggc acg gag act ggt aat gcg gag gcg
 S   M   G   N   A   L   I   L   Y   G   T   E   T   G   N   A   E   A
    CinC gene starts
tgt gcg acc aca att tcg cag gtt ctg gcg gat acc gtt gat acg aag gtg cat
 C   A   T   T   I   S   Q   V   L   A   D   T   V   D   T   K   V   H

gat ctg gcc gat atg acc ccc cgc gca atg ctg gat agt ggt gcg gat ctg atc
 D   L   A   D   M   T   P   R   A   M   L   D   S   G   A   D   L   I

gtg ttt gcc acg gca acc tat ggc gaa ggc gaa ttt gcg ggt ggt ggc gca gcg
 V   F   A   T   A   T   Y   G   E   G   E   F   A   G   G   G   A   A

ttc ttt gag act ctg cgt gaa acc aaa ccc gat ctg agc ggt ctg cgt ttc gct
 F   F   E   T   L   R   E   T   K   P   D   L   S   G   L   R   F   A

gtg ttt ggt ctg ggc gat agc tat tac act acg ttt aac cag gcc ggt gca act
 V   F   G   L   G   D   S   Y   Y   T   T   F   N   Q   A   G   A   T

gcc gca acc att tta gcc agc ctg ggt ggt acc cag gtg ggt gat acg gcg cgc
 A   A   T   I   L   A   S   L   G   G   T   Q   V   G   D   T   A   R

cac gat act tcg tcg ggc gac gat cca gaa gaa acc gcc gaa gaa tgg gcc cgc
 H   D   T   S   S   G   D   D   P   E   E   T   A   E   E   W   A   R

gag atc ttg act gca ttg gca act ccg gca gtt agt taa taa
 E   I   L   T   A   L   A   T   P   A   V   S   -   -
                                              Stop codon
```

7. REFERENCES

Aharoni, A., Griffiths, A.D., Tawfik, D.S., (2005) High-throughput screens and selections of enzyme-encoding genes. Current opinion in Chemical Biology 9, 210-216.

Aliverti, A., Curti, B., Vanoni, M.A., (1999) Identifying and quantitating FAD and FMN in simple and in iron-sulfur-containing flavoproteins. Flavoprotein protocols. Methods in Molecular Biology. Springer 131, 9-23.

Andersson, M., Holmberg, H., Adlercreutz, P., (1997) Evaluation of electron mediators for the electromicrobial reduction of chloropyruvate by *Proteus vulgaris* cells. Enzyme and Microbial Technology 20, 150-156.

Arai, R., Ueda, H., Kitayama, A., Kamiya, N., Nagamune, T., (2001) Design of the linkers which effectively separate domains of a bifunctional fusion protein. Protein Engineering 14, 529-532.

Arnold, F.H., (2001) Combinatorial and computational challenges for biocatalyst design. Nature 409, 253-257.

Arnold, F.H., Georgiou, G., (2003) Directed evolution library creation. Methods in Molecular Biology Humana Press 231.

Badalassi, F., Wahler, D., Klein, G., Crotti, P., Reymond, J.L., (2000) A Versatile Periodate-Coupled Fluorogenic Assay for Hydrolytic Enzymes. Angewandte Chemie 112, 4233-4236.

Bao, L., Sun, D., Tachikawa, H., Davidson, V.L., (2002) Improved sensitivity of a histamine sensor using an engineered methylamine dehydrogenase. Analytical Chemistry 74, 1144-1148.

Belsare, K.D., Ruff, A.J., Martinez, R., Shivange, A.V., Mundhada, H., Holtmann, D., Schrader, J., Schwaneberg, U., (2014) P-LinK: A method for generating multicomponent cytochrome P450 fusions with variable linker length. Biotechniques 57, 13-20.

Bernhardt, R., (2006) Cytochromes P450 as versatile biocatalysts. Journal of Biotechnology 124, 128-145.

Bizzarri, A., Cannistraro, S., (2005) Electron transfer in metalloproteins. Encyclopedia of Condensed Matter Physics. Elsevier, Dordrecht, The Netherlands.

Blanusa, M., Schenk, A., Sadeghi, H., Marienhagen, J., Schwaneberg, U., (2010) Phosphorothioate-based ligase-independent gene cloning (PLICing): an enzyme-free and sequence-independent cloning method. Analytical Biochemistry 406, 141-146.

Bornscheuer, U., Huisman, G., Kazlauskas, R., Lutz, S., Moore, J., Robins, K., (2012) Engineering the third wave of biocatalysis. Nature 485, 185-194.

Cadwell, R.C., Joyce, G.F., (1994) Mutagenic PCR. Genome Research 3, S136-S140.

Carmichael, A.B., Wong, L.L., (2001) Protein engineering of Bacillus megaterium CYP102. European Journal of Biochemistry 268, 3117-3125.

Çekiç, S.Z., Holtmann, D., Güven, G., Mangold, K.-M., Schwaneberg, U., Schrader, J., (2010) Mediated electron transfer with P450cin. Electrochemistry Communications 12, 1547-1550.

Çelik, A., Speight, R.E., Turner, N.J., (2005) Identification of broad specificity P450 CAM variants by primary screening against indole as substrate. Chemical Communications, 3652-3654.

Chica, R.A., Doucet, N., Pelletier, J.N., (2005) Semi-rational approaches to engineering enzyme activity: combining the benefits of directed evolution and rational design. Current Opinion in Biotechnology 16, 378-384.

Chou, P.Y., Fasman, G.D., (1974a) Conformational parameters for amino acids in helical, β-sheet, and random coil regions calculated from proteins. Biochemistry 13, 211-222.

Chou, P.Y., Fasman, G.D., (1974b) Prediction of protein conformation. Biochemistry 13, 222-245.

Cirino, P.C., Arnold, F.H., (2003) A Self-Sufficient Peroxide-Driven Hydroxylation Biocatalyst. Angewandte Chemie 115, 3421-3423.

Cooper, G.M., (2000) The Central Role of Enzymes as Biological Catalysts In: The Cell: A Molecular Approach. Sinauer Associates, Sunderland.

Cordes, M., Giese, B., (2009) Electron transfer in peptides and proteins. Chemical Society Reviews 38, 892-901.

Cox, B., Parry, J., (1968) The isolation, genetics and survival characteristics of ultraviolet light-sensitive mutants in yeast. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis 6, 37-55.

Cryle, M.J., Stok, J.E., De Voss, J.J., (2003) Reactions catalyzed by bacterial cytochromes P450. Australian Journal of Chemistry 56, 749-762.

Datta, S., Holmes, B., Park, J.I., Chen, Z., Dibble, D.C., Hadi, M., Blanch, H.W., Simmons, B.A., Sapra, R., (2010) Ionic liquid tolerant hyperthermophilic cellulases for biomass pretreatment and hydrolysis. Green Chemistry 12, 338-345.

Davidson, V.L., (2003) Probing mechanisms of catalysis and electron transfer by methylamine dehydrogenase by site-directed mutagenesis of αPhe[55]. Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics 1647, 230-233.

De Montellano, P.R.O., (2005) Cytochrome P450: structure, mechanism, and biochemistry. Springer Edition 3.

Denisov, I.G., Makris, T.M., Sligar, S.G., Schlichting, I., (2005) Structure and chemistry of cytochrome P450. Chemical Reviews 105, 2253-2278.

Dennig, A., Marienhagen, J., Ruff, A.J., Guddat, L., Schwaneberg, U., (2012) Directed Evolution of P 450 BM 3 into *p*-Xylene Hydroxylase. ChemCatChem 4, 771-773.

Eckstein, F., Gish, G., (1989) Phosphorothioates in molecular biology. Trends in biochemical sciences 14, 97-100.

Farinas, E.T., (2006) Fluorescence activated cell sorting for enzymatic activity. Combinatorial Chemistry & High Throughput Screening 9, 321-328.

Farinas, E.T., Bulter, T., Arnold, F.H., (2001) Directed enzyme evolution. Current Opinion in Biotechnology 12, 545-551.

Fasan, R., (2012) Tuning P450 Enzymes as Oxidation Catalysts. ACS Catalysis 2, 647-666.

Faulkner, K.M., Shet, M.S., Fisher, C.W., Estabrook, R.W., (1995) Electrocatalytically driven omega-hydroxylation of fatty acids using cytochrome P450 4A1. Proceedings of the National Academy of Sciences 92, 7705-7709.

Ferapontova, E., Schmengler, K., Börchers, T., Ruzgas, T., Gorton, L., (2002) Effect of cysteine mutations on direct electron transfer of horseradish peroxidase on gold. Biosensors and Bioelectronics 17, 953-963.

Fleming, B.D., Johnson, D.L., Bond, A.M., Martin, L.L., (2006) Recent progress in cytochrome P450 enzyme electrochemistry 2, 581-589.

Geier, M., Braun, A., Fladischer, P., Stepniak, P., Rudroff, F., Hametner, C., Mihovilovic, M.D., Glieder, A., (2013) Double site saturation mutagenesis of the human cytochrome P450 2D6 results in regioselective steroid hydroxylation. FEBS Journal 280, 3094-3108.

George, R.A., Heringa, J., (2002) An analysis of protein domain linkers: their classification and role in protein folding. Protein Engineering 15, 871-879.

Gillam, E.M., Notley, L.M., Cai, H., De Voss, J.J., Guengerich, F.P., (2000) Oxidation of indole by cytochrome P450 enzymes. Biochemistry 39, 13817-13824.

Gish, G., Eckstein, F., (1988) DNA and RNA sequence determination based on phosphorothioate chemistry. Science 240, 1520-1522.

Glieder, A., Meinhold, P., (2003) High-throughput screens based on NAD(P)H depletion. Directed Enzyme Evolution. Methods in Molecular Biology Springer 230, 157-170.

Goodsell, D.S., Morris, G.M., Olson, A.J., (1996) Automated docking of flexible ligands: applications of AutoDock. Journal of Molecular Recognition 9, 1-5.

Govindaraj, S., Poulos, T.L., (1995) Role of the linker region connecting the reductase and heme domains in cytochrome P450BM-3. Biochemistry 34, 11221-11226.

Govindaraj, S., Poulos, T.L., (1996) Probing the structure of the linker connecting the reductase and heme domains of cytochrome P450BM-3 using site-directed mutagenesis. Protein science 5, 1389-1393.

Gray, H.B., Winkler, J.R., (2009) Electron flow through proteins. Chemical Physics Letters 483, 1-9.

Gray, H.B., Winkler, J.R., (2010) Electron flow through metalloproteins. Biochimica et Biophysica Acta (BBA)-Bioenergetics 1797, 1563-1572.

Guengerich, F.P., (1995) Cytochrome P450 proteins and potential utilization in biodegradation. Environmental Health Perspectives 103, 25-28.

Gupta, R., Beg, Q., Lorenz, P., (2002) Bacterial alkaline proteases: molecular approaches and industrial applications. Applied Microbiology and Biotechnology 59, 15-32.

Güven, G., Prodanovic, R., Schwaneberg, U., (2010) Protein engineering–an option for enzymatic biofuel cell design. Electroanalysis 22, 765-775.

Hannemann, F., Bichet, A., Ewen, K.M., Bernhardt, R., (2007) Cytochrome P450 systems—biological variations of electron transport chains. Biochimica et Biophysica Acta (BBA)-General Subjects 1770, 330-344.

Hawkes, D.B., Adams, G.W., Burlingame, A.L., de Montellano, P.R.O., De Voss, J.J., (2002) Cytochrome P450cin (CYP176A), isolation, expression, and characterization. Journal of Biological Chemistry 277, 27725-27732.

Hawkes, D.B., Slessor, K.E., Bernhardt, P.V., De Voss, J.J., (2010) Cloning, Expression and Purification of Cindoxin, an Unusual Fmn-Containing Cytochrome P450 Redox Partner. ChemBioChem 11, 1107-1114.

Hendricks, C.L., Ross, J.R., Pichersky, E., Noel, J.P., Zhou, Z.S., (2004) An enzyme-coupled colorimetric assay for *S*-adenosylmethionine-dependent methyltransferases. Analytical Biochemistry 326, 100-105.

Hirakawa, H., Kamiya, N., Tanaka, T., Nagamune, T., (2007) Intramolecular electron transfer in a cytochrome P450cam system with a site-specific branched structure. Protein Engineering Design and Selection 20, 453-459.

Hirakawa, H., Nagamune, T., (2010) Molecular assembly of P450 with ferredoxin and ferredoxin reductase by fusion to PCNA. ChemBioChem 11, 1517-1520.

Hollmann, F., Hofstetter, K., Schmid, A., (2006) Non-enzymatic regeneration of nicotinamide and flavin cofactors for monooxygenase catalysis. Trends in Biotechnology 24, 163-171.

Holtmann, D., Schrader, J., (2007) Approaches to Recycling and Substituting NAD (P) H as a CYP Cofactor. Modern Biooxidation: Enzymes, Reactions and Applications, Wiley- VCH Verlag GmbH & Co. 265-290.

Ichikawa, Y., Yamano, T., (1967) Reconversion of detergent-and sulfhydryl reagent-produced P-420 to P-450 by polyols and glutathione. Biochimica et Biophysica Acta (BBA)-Bioenergetics 131, 490-497.

Jenkins, C.M., Waterman, M.R., (1994) Flavodoxin and NADPH-flavodoxin reductase from Escherichia coli support bovine cytochrome P450c17 hydroxylase activities. Journal of Biological Chemistry 269, 27401-27408.

Joo, H., Arisawa, A., Lin, Z., Arnold, F.H., (1999a) A high-throughput digital imaging screen for the discovery and directed evolution of oxygenases. Chemistry & Biology 6, 699-706.

Joo, H., Lin, Z., Arnold, F.H., (1999b) Laboratory evolution of peroxide-mediated cytochrome P450 hydroxylation. Nature 399, 670-673.

Jung, S.T., Lauchli, R., Arnold, F.H., (2011) Cytochrome P450: taming a wild type enzyme. Current Opinion in Biotechnology 22, 809-817.

Kardashliev, T., Ruff, A.J., Zhao, J., Schwaneberg, U., (2013) A High-Throughput Screening Method to Reengineer DNA Polymerases for Random Mutagenesis. Molecular Biotechnology 3, 274-283.

Kellner, D.G., Maves, S.A., Sligar, S.G., (1997) Engineering cytochrome P450s for bioremediation. Current Opinion in Biotechnology 8, 274-278.

Kim, D., de Montellano, P.R.O., (2009) Tricistronic overexpression of cytochrome P450 cam, putidaredoxin, and putidaredoxin reductase provides a useful cell-based catalytic system. Biotechnology Letters 31, 1427-1431.

Kimmich, N., Das, A., Sevrioukova, I., Meharenna, Y., Sligar, S.G., Poulos, T.L., (2007) Electron transfer between cytochrome P450cin and its FMN-containing redox partner, cindoxin. Journal of Biological Chemistry 282, 27006-27011.

Kitazume, T., Haines, D.C., Estabrook, R.W., Chen, B., Peterson, J.A., (2007) Obligatory intermolecular electron-transfer from FAD to FMN in dimeric P450BM-3. Biochemistry 46, 11892-11901.

Koeller, K.M., Wong, C.-H., (2001) Enzymes for chemical synthesis. Nature 409, 232-240.

Korkegian, A., Black, M.E., Baker, D., Stoddard, B.L., (2005) Computational thermostabilization of an enzyme. Science 308, 857-860.

Krieger, E., Koraimann, G., Vriend, G., (2002) Increasing the precision of comparative models with YASARA NOVA—a self-parameterizing force field. Proteins: Structure, Function, and Bioinformatics 47, 393-402.

Kubo, T., Peters, M.W., Meinhold, P., Arnold, F.H., (2006) Enantioselective Epoxidation of Terminal Alkenes to (R)-and (S)-Epoxides by Engineered Cytochromes P450 BM-3. Chemistry-A European Journal 12, 1216-1220.

Kuhad, R.C., Gupta, R., Singh, A., (2011) Microbial cellulases and their industrial applications. Enzyme Research 1-10.

Leemhuis, H., Kelly, R.M., Dijkhuizen, L., (2009) Directed evolution of enzymes: library screening strategies. IUBMB life 61, 222-228.

Lehmann, C., Sibilla, F., Maugeri, Z., Streit, W.R., de María, P.D., Martinez, R., Schwaneberg, U., (2012) Reengineering CelA2 cellulase for hydrolysis in aqueous solutions of deep eutectic solvents and concentrated seawater. Green Chemistry 14, 2719-2726.

Ley, C., Schewe, H., Ströhle, F.W., Ruff, A.J., Schwaneberg, U., Schrader, J., Holtmann, D., (2013) Coupling of electrochemical and optical measurements in a microtiter plate for the fast development of electro enzymatic processes with P450s. Journal of Molecular Catalysis B: Enzymatic 92, 71-78.

Lin, C.-L., Tang, Y.-L., Lin, S.-M., (2011) Efficient bioconversion of compactin to pravastatin by the quinoline-degrading microorganism *Pseudonocardia carboxydivorans* isolated from petroleum-contaminated soil. Bioresource Technology 102, 10187-10193.

Lin, H., Cornish, V.W., (2002) Screening and selection methods for large-scale analysis of protein function. Angewandte Chemie International Edition 41, 4402-4425.

Madrona, Y., Hollingsworth, S.A., Tripathi, S., Fields, J.B., Rwigema, J.-C.N., Tobias, D.J., Poulos, T.L., (2014) Crystal Structure of Cindoxin, the P450cin Redox Partner. Biochemistry 53, 1435-1446.

Makris, T.M., Denisov, I., Schlichting, I., Sligar, S.G., (2005) Activation of molecular oxygen by cytochrome P450. Cytochrome P450. Springer, 149-182.

Marcus, R.A., Sutin, N., (1985) Electron transfers in chemistry and biology. Biochimica et Biophysica Acta (BBA)-Reviews on Bioenergetics 811, 265-322.

Martinez, R., Jakob, F., Tu, R., Siegert, P., Maurer, K.H., Schwaneberg, U., (2013) Increasing activity and thermal resistance of Bacillus gibsonii alkaline protease (BgAP) by directed evolution. Biotechnology and Bioengineering 110, 711-720.

Maurer, K.-H., (2004) Detergent proteases. Current Opinion in Biotechnology 15, 330-334.

Maurer, S.C., Kühnel, K., Kaysser, L.A., Eiben, S., Schmid, R.D., Urlacher, V.B., (2005) Catalytic hydroxylation in biphasic systems using CYP102A1 mutants. Advanced Synthesis & Catalysis 347, 1090-1098.

Maurer, S.C., Schulze, H., Schmid, R.D., Urlacher, V., (2003) Immobilisation of P450 BM-3 and an NADP+ Cofactor Recycling System: Towards a Technical Application of Heme-Containing Monooxygenases in Fine Chemical Synthesis. Advanced Synthesis & Catalysis 345, 802-810.

McLean, K., Sabri, M., Marshall, K., Lawson, R., Lewis, D., Clift, D., Balding, P., Dunford, A., Warman, A., McVey, J., (2005) Biodiversity of cytochrome P450 redox systems. Biochemical Society Transactions 33, 796-801.

McLean, K.J., Girvan, H.M., Munro, A.W., (2007) Cytochrome P450/redox partner fusion enzymes: biotechnological and toxicological prospects. Current Opinion on Drug Metabolism and Toxicology 3, 847-863.

Meharenna, Y.T., Li, H., Hawkes, D.B., Pearson, A.G., De Voss, J., Poulos, T.L., (2004) Crystal structure of P450cin in a complex with its substrate, 1, 8-cineole, a close structural homologue to D-camphor, the substrate for P450cam. Biochemistry 43, 9487-9494.

Meunier, B., De Visser, S.P., Shaik, S., (2004) Mechanism of oxidation reactions catalyzed by cytochrome P450 enzymes. Chemical Reviews 104, 3947-3980.

Michizoe, J., Ichinose, H., Kamiya, N., Maruyama, T., Goto, M., (2005) Functionalization of the cytochrome P450cam monooxygenase system in the cell-like aqueous compartments of water-in-oil emulsions. Journal of Bioscience and Bioengineering 99, 12-17.

Miller, O.J., Bernath, K., Agresti, J.J., Amitai, G., Kelly, B.T., Mastrobattista, E., Taly, V., Magdassi, S., Tawfik, D.S., Griffiths, A.D., (2006) Directed evolution by in vitro compartmentalization. Nature Methods 3, 561-570.

Miyazawa, M., Hashimoto, Y., (2002) Antimicrobial and bactericidal activities of esters of 2-endo-hydroxy-1, 8-cineole as new aroma chemicals. Journal of Agricultural and Food Chemistry 50, 3522-3526.

Miyazawa, M., Shindo, M., Shimada, T., (2001) Oxidation of 1, 8-cineole, the monoterpene cyclic ether originated from Eucalyptus polybractea, by cytochrome P450 3A enzymes in rat and human liver microsomes. Drug Metabolism and Disposition 29, 200-205.

Mohan, U., Banerjee, U.C., (2008) Molecular evolution of a defined DNA sequence with accumulation of mutations in a single round by a dual approach to random chemical mutagenesis (DuARCheM). ChemBioChem 9, 2238-2243.

Müller, C.A., Akkapurathu, B., Winkler, T., Staudt, S., Hummel, W., Gröger, H., Schwaneberg, U., (2013) In Vitro Double Oxidation of n-Heptane with Direct Cofactor Regeneration. Advanced Synthesis & Catalysis 355, 1787-1798.

Mundhada, H., Marienhagen, J., Scacioc, A., Schenk, A., Roccatano, D., Schwaneberg, U., (2011) SeSaM-Tv-II Generates a Protein Sequence Space that is Unobtainable by epPCR. ChemBioChem 12, 1595-1601.

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.

Murakami, K., Mason, H., (1967) An electron spin resonance study of microsomal Fex. Journal of Biological Chemistry 242, 1102-1110.

Nazor, J., (2007) Alternative Cofactor Systems Driving P450 BM3. Jacobs University Bremen.

Nazor, J., Dannenmann, S., Adjei, R.O., Fordjour, Y.B., Ghampson, I.T., Blanusa, M., Roccatano, D., Schwaneberg, U., (2008) Laboratory evolution of P450 BM3 for mediated electron transfer yielding an activity-improved and reductase-independent variant. Protein Engineering Design and Selection 21, 29-35.

Nazor, J., Schwaneberg, U., (2006) Laboratory Evolution of P450 BM-3 for Mediated Electron Transfer. ChemBioChem 7, 638-644.

Olsen, M., Iverson, B., Georgiou, G., (2000) High-throughput screening of enzyme libraries. Current Opinion in Biotechnology 11, 331-337.

Omura, T., Sato, R., (1964) The carbon monoxide-binding pigment of liver microsomes. I. Evidence for its hemoprotein nature. Journal of Biological Chemistry 239, 2370-2378.

Petzoldt K, A.K., Laurent H, Wiechert R, (1982) Process for the preparation of 11-beta-hydroxy steroids. US Patent, USA.

Poulos, T.L., Finzel, B.C., Howard, A.J., (1987) High-resolution crystal structure of cytochrome P450cam. Journal of Molecular Biology 195, 687-700.

Prasad, S., Murugan, R., Mitra, S., (2005) An artificial electron donor supported catalytic cycle of *Pseudomonas putida* cytochrome $P450_{cam}$. Biochemical and Biophysical Research Communications 335, 590-595.

Reetz, M.T., Jaeger, K.E., (2000) Enantioselective enzymes for organic synthesis created by directed evolution. Chemistry-A-European Journal 6, 407-412.

Reetz, M.T., Kahakeaw, D., Lohmer, R., (2008) Addressing the numbers problem in directed evolution. ChemBioChem 9, 1797-1804.

Reipa, V., Mayhew, M.P., Vilker, V.L., (1997) A direct electrode-driven P450 cycle for biocatalysis. Proceedings of the National Academy of Sciences 94, 13554-13558.

Robin, A., Roberts, G.A., Kisch, J., Sabbadin, F., Grogan, G., Bruce, N., Turner, N.J., Flitsch, S.L., (2009) Engineering and improvement of the efficiency of a chimeric [P450cam-RhFRed reductase domain] enzyme. Chemical Communications, 2478-2480.

Rodríguez, P., Sierra, W., Rodríguez, S., Menéndez, P., (2006) Biotransformation of 1, 8-cineole, the main product of Eucalyptus oils. Electronic Journal of Biotechnology 9.

Ruff, A.J., (2012) Advances in directed monooxygenase evolution: from diversity generation and flow cytometry screening to tailor-made monooxygenases. Universitätsbibliothek.

Ruff, A.J.l., Dennig, A., Wirtz, G., Blanusa, M., Schwaneberg, U., (2012) Flow cytometer-based high-throughput screening system for accelerated directed evolution of P450 monooxygenases. ACS Catalysis 2, 2724-2728.

Saab-Rincón, G., Valderrama, B., (2009) Protein engineering of redox-active enzymes. Antioxidants & Redox Signaling 11, 167-192.

Sabbadin, F., Hyde, R., Robin, A., Hilgarth, E.M., Delenne, M., Flitsch, S., Turner, N., Grogan, G., Bruce, N.C., (2010) LICRED: A Versatile Drop-In Vector for Rapid Generation of Redox-Self-Sufficient Cytochrome P450s. ChemBioChem 11, 987-994.

Sambrook, J., Fritsch, E.F., Maniatis, T., (1989) Molecular Cloning. Cold Spring Harbor Laboratory Press New York.

Sambrook, J.R., Russell, D.W., (2001) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press New York.

Schmid, A., Dordick, J., Hauer, B., Kiener, A., Wubbolts, M., Witholt, B., (2001) Industrial biocatalysis today and tomorrow. Nature 409, 258-268.

Schmidt-Dannert, C., Arnold, F.H., (1999) Directed evolution of industrial enzymes. Trends in Biotechnology 17, 135-136.

Schwaneberg, U., Appel, D., Schmitt, J., Schmid, R.D., (2000) P450 in biotechnology: zinc driven ω-hydroxylation of p-nitrophenoxydodecanoic acid using P450 BM-3 F87A as a catalyst. Journal of Biotechnology 84, 249-257.

Sevrioukova, I.F., Li, H., Zhang, H., Peterson, J.A., Poulos, T.L., (1999) Structure of a cytochrome P450–redox partner electron-transfer complex. Proceedings of the National Academy of Sciences 96, 1863-1868.

Shumyantseva, V., Bulko, T., Rudakov, Y.O., Kuznetsova, G., Samenkova, N., Lisitsa, A., Karuzina, I., Archakov, A., (2007) Electrochemical properties of cytochroms P450 using nanostructured electrodes: direct electron transfer and electro catalysis. Journal of Inorganic Biochemistry 101, 859-865.

Shumyantseva, V.V., Bulko, T.V., Archakov, A.I., (2005) Electrochemical reduction of cytochrome P450 as an approach to the construction of biosensors and bioreactors. Journal of Inorganic Biochemistry 99, 1051-1063.

Shumyantseva, V.V., Bulko, T.V., Bachmann, T.T., Bilitewski, U., Schmid, R.D., Archakov, A.I., (2000) Electrochemical reduction of flavocytochromes 2B4 and 1A2 and their catalytic activity. Archives of Biochemistry and Biophysics 377, 43-48.

Shumyantseva, V.V., Bulko, T.V., Schmid, R.D., Archakov, A.I., (2002) Photochemical properties of a riboflavins/cytochrome P450 2B4 complex. Biosensors and Bioelectronics 17, 233-238.

Sibbesen, O., De Voss, J.J., de Montellano, P.R.O., (1996) Putidaredoxin reductase-putidaredoxin-cytochrome P450cam triple fusion protein construction of a self-sufficient Escherichia coli catalytic system. Journal of Biological Chemistry 271, 22462-22469.

Slessor, K.E., Hawkes, D.B., Farlow, A., Pearson, A.G., Stok, J.E., De Voss, J.J., (2012) An *in vivo* cytochrome P450$_{cin}$ (CYP176A1) catalytic system for metabolite production. Journal of Molecular Catalysis B: Enzymatic 79, 15-20.

Slessor, K.E., Stok, J.E., Cavaignac, S.M., Hawkes, D.B., Ghasemi, Y., De Voss, J.J., (2010) Cineole biodegradation: Molecular cloning, expression and characterisation of (1*R*)-6β-hydroxycineole dehydrogenase from *Citrobacter braakii*. Bioorganic Chemistry 38, 81-86.

Stemmer, W., Holland, B., (2003) Survival of the fittest molecule. American Scientist 91, 526-533.

Stemmer, W.P., (1994) Rapid evolution of a protein in vitro by DNA shuffling. Nature 370, 389-391.

Tishkov, V.I., Galkin, A.G., Fedorchuk, V.V., Savitsky, P.A., Rojkova, A.M., Gieren, H., Kula, M.-R., (1999) Pilot scale production and isolation of recombinant NAD+-and NADP+-specific formate dehydrogenases. Biotechnology and bioengineering 64, 187-193.

Tosstorff, A., Dennig, A., Ruff, A.J., Schwaneberg, U., Sieber, V., Mangold, K.-M., Schrader, J., Holtmann, D., (2014) Mediated electron transfer with monooxygenases–insight in interactions between reduced mediators and the co-substrate oxygen. Journal of Molecular Catalysis B: Enzymatic 108, 51-58.

Udit, A.K., Gray, H.B., (2005) Electrochemistry of heme–thiolate proteins. Biochemical and Biophysical Research Communications 338, 470-476.

Urban, A., Neukirchen, S., Jaeger, K.-E., (1997) A rapid and efficient method for site-directed mutagenesis using one-step overlap extension PCR. Nucleic Acids Research 25, 2227-2228.

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., 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.

Wandrey, C., (2004) Biochemical reaction engineering for redox reactions. The Chemical Record 4, 254-265.

Whitehouse, C.J., Bell, S.G., Wong, L.-L., (2012) P450BM3 (CYP102A1): connecting the dots. Chemical Society Reviews 41, 1218-1260.

Wichmann, R., Vasic-Racki, D., (2005) Cofactor regeneration at the lab scale. Technology Transfer in Biotechnology. Springer 92, 225-260.

Wong, T.S., Roccatano, D., Schwaneberg, U., (2007) Steering directed protein evolution: strategies to manage combinatorial complexity of mutant libraries. Environmental Microbiology 9, 2645-2659.

Wong, T.S., Schwaneberg, U., (2003) Protein engineering in bioelectrocatalysis. Current Opinion in Biotechnology 14, 590-596.

Wong, T.S., Tee, K.L., Hauer, B., Schwaneberg, U., (2004) Sequence saturation mutagenesis (SeSaM): a novel method for directed evolution. Nucleic Acids Research 32, e26.

Woodley, J.M., (2008) New opportunities for biocatalysis: making pharmaceutical processes greener. Trends in Biotechnology 26, 321-327.

Wriggers, W., Chakravarty, S., Jennings, P.A., (2005) Control of protein functional dynamics by peptide linkers. Peptide Science 80, 736-746.

You, L., Arnold, F., (1996) Directed evolution of subtilisin E in Bacillus subtilis to enhance total activity in aqueous dimethylformamide. Protein Engineering 9, 77-83.

Zhao, H., Giver, L., Shao, Z., Affholter, J.A., Arnold, F.H., (1998) Molecular evolution by staggered extension. Nature Biotechnology 16, 258-261.

Zhao, Q., Modi, S., Smith, G., Paine, M., McDonagh, P.D., Wolf, C.R., Tew, D., Lian, L.-Y., Roberts, G.C., Driessen, H.P., (1999) Crystal structure of the FMN-binding domain of human cytochrome P450 reductase at 1.93 Å resolution. Protein Science 8, 298-306.

Personal data

Name: Ketaki D. Belsare
Gender: Female
Date of birth: 13th December, 1986
Place of birth: Nagpur, India
Nationality: Indian

Education

2011 – Present   **RWTH Aachen University, Aachen, Germany**
                   PhD student
                   Thesis: "Directed Evolution of Cytochrome P450cin for Mediated Electron Transfer"

2008 – 2010   **M. Tech in Pharmaceutical Biotechnology**
                   National Institute of Pharmaceutical Education and Research, India.

2004 – 2008   **Bachelor of Pharmacy**
                   University of Pune, India

2002 - 2004   **Higher Secondary Examination**, Nasik, India.

1992 - 2002   **Secondary School Examination**, Nasik, India.

_____

Ketaki D. Belsare, Aachen

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe, dass alle Stellen der Arbeit, die wörtlich oder sinngemäß aus anderen Quellen übernommen wurden, als solche kenntlich gemacht sind und dass die Arbeit in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegt wurde.

Aachen, November 2014

Ketaki Deepak Belsare