Brigitte Mathilde Höller

New Oxidoreductases for Biocatalysis

DISSERTATION

zur Erlangung des akademischen Grades einer Doktorin der Naturwissenschaften

erreicht an der

Technischen Universität Graz

in Zusammenarbeit mit

Ao. Univ. Prof. Mag.rer.nat Dr.rer.nat Anton Glieder

Institut für Molekulare Biotechnologie

Technische Universität Graz

2010

Danksagung

Ich möchte mich an dieser Stelle bei allen Personen bedanken, die zum Gelingen dieser Arbeit in unterschiedlichster Weise beigetragen haben.

Allen voran möchte ich mich bei Ao. Univ. Prof. Anton Glieder für die Betreuung der Dissertation herzlich bedanken. Danke für die gewährte Freiheit bei der Durchführung der Arbeit und insbesondere für die wissenschaftlichen Ratschläge.

Besonderer Dank gilt auch Dr. Oliver May, Dr. Stefaan De Wildeman, Dr. Natascha Sereinig und Dr. Martin Schürmann von DSM, sowohl für die finanzielle Unterstützung als auch für die angenehme Atmosphäre während der zahlreichen Meetings.

Ich möchte mich für die wertvollen Gespräche und wissenschaftliche Unterstützung sowie für geniale Lachmomente und einzigartige Erinnerungen, einerseits bei meinen „früheren" Kolleg(inn)en Franz Hartner, Roland Weis, Matthias Schittmayer, Beate Pscheidt, Alexander Kern, Bianca Hecher, Kerstin Kitz, Bettina Janesch, Christian Gruber und Thomas Züllig, und andererseits bei der „neuen" Truppe bedanken. Hierbei vor allem bei meinen lieben Bürokolleg(inn)en Margit Winkler, Barbara Petschacher und Karlheinz Flicker.

In besonderem Maße zur guten Atmosphäre sowohl im Labor als auch bei vielen, unvergesslichen Freizeitaktivitäten beigetragen haben: Uli Schreiner, Hanni Mandl und Mandy Vink! Danke Mädels für eure Freundschaft!

Darüber hinaus möchte ich mich bei allen Personen in anderen Arbeitsgruppen bzw. Instituten, sowie bei Thomas Bergner und Gerald Mascher für ihre gute Mitarbeit im Ene-Reduktase Projekt, bedanken.

Mein ganz besonderer Dank gilt meinen Eltern, denen diese Dissertation gewidmet ist. Ihr habt mir mein Studium ermöglicht und immer an mich und meine Fähigkeiten geglaubt. Danke für eure Liebe!

Ein großes Dankeschön geht an meinen Freund Christian, der mich im letzten Jahr meiner Dissertation bestens unterstützt hat. Danke für dein Verständnis in manch schweren Momenten und die vielen lieben und aufmunternden Worte, welche immer wieder zu meiner Motivation beigetragen haben. Schön, dass es dich gibt!

Für meine Eltern

Deutsche Fassung:

Beschluss der Curricula-Kommission für Bachelor-, Master- und Diplomstudien vom 10.11.2008

Genehmigung des Senates am 1.12.2008

EIDESSTATTLICHE ERKLÄRUNG

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommene Stellen als solche kenntlich gemacht habe.

Graz, am …… ……..

(Unterschrift)

Englische Fassung:

STATUTORY DECLARATION

I declare that I have authored this thesis independently, that I have not used other than the declared sources / resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

…… ……..

date (signature)

Abstract

New Oxidoreductases for Biocatalysis

Protein engineering is an indispensable tool in the field of applied biocatalysis and new methods for identification of enzymes, enzyme engineering and high throughput analytics are a constant demand. In the first part of this work we describe the Transposon Integration mediated Mutagenesis (TIM) which is a broadly applicable tool for protein engineering. TIM enables both insertions and deletions of non coding sequences. This method combines random integration of modified bacteriophage *Mu* transposons with their subsequent defined excision employing type IIS restriction endonuclease *Aar*I. The employed strategy facilitates the isolation of randomly tagged active enzymes (e.g. 6x His Tag) or insertion and deletion of other sequences with a defined number of bases in single mutagenesis experiments.

In the second part of this study we were pursuing the goal to engineer enzymes from the industrially relevant enzyme class of non-heme oxidoreductases. As models for this work we have chosen an interesting amine dehydrogenase from *Streptomyces virginiae* and an 'ene'-reductase. This study describes the heterologous expression in *Escherichia coli*, the establishment of a screening system and the development of analytical methods (TLC, GC and HPLC). Several rounds of error-prone PCR were employed for activity improvement towards new substrates in the presence of 2% of an organic co-solvent (DMSO, DMF or NMP). A SlonoMax® library with two interesting amino acid positions saturated and a semi-rational approach using active-site saturation was performed to further enhance the enzyme activity. The best variant showed a significant increase in activity in the presence of 2% DMSO and 2% DMF and an even higher improvement in the presence of 2% NMP. The selectivity of the enzyme was retained.

Zusammenfassung

Neue Oxidoreduktasen für die Biokatalyse

Proteinentwicklung ist unverzichtbar im Bereich der angewandten Biokatalyse und neue Methoden für die Enzymidentifizierung, Enzymentwicklung- und Hochdurchsatzanalytik sind ständig von Interesse. Der erste Teil dieser Dissertation beschreibt eine Transposon Mutagenese Methode (TIM), welche in der Proteinentwicklung wie auch zur Insertions/Deletionsmutagenese nicht kodierender Nukleinsäure Sequenzen breit anwendbar ist. Diese Methode verbindet die zufällige Integration von veränderten Transposons vom Bakteriophagen *Mu* mit einem definierten Ausschneiden aus der DNA mit Hilfe der Typ IIS Restriktionsendonuklease *Aar*I. Mithilfe der angewandten Strategie ist es möglich, in einem einzigen Mutageneseexperiment Enzyme mit integrierten Tags (z.B. 6x His Tag) oder anderen Insertionssequenzen bzw. Deletionen definierter Größe an verschiedenen und zufälligen Positionen herzustellen.

Im zweiten Teil dieser Studie wurde das Ziel der Verbesserung von Enzymen der industriell wichtigen Enzymklasse der häm-unabhängigen-Oxidoreduktasen verfolgt. Als Modelle für diese Studien wurden eine Amindehydrogenase von *Streptomyces virginiae* und eine 'ene'–Reduktase gewählt. Diese Arbeit beschreibt sowohl heterologe Expression in *Escherichia coli,* die Etablierung eines Screeningsystems, inklusive Entwicklung analytischer Methoden wie DC, GC und HPLC. Mehrere Runden von fehleranfälliger PCR wurden angewandt, um die Aktivität des Enzyms für neue Substrate in Anwesenheit von 2% eines organischen Lösungsmittels (DMSO, DMF oder NMP) zu verbessern. Eine SlonoMax® Genbank (zwei interessante Aminosäurepositionen gesättigt) wie auch ein semi-rationaler Ansatz zur Sättigung von interessanten Positionen im aktiven Zentrum wurden verwendet, um die Enzymaktivität weiter zu erhöhen. Die beste Mutante zeigte eine mehrfache Steigerung der Aktivität in Anwesenheit von 2% DMSO und 2% DMF, sowie eine noch höhere Verbesserung in Anwesenheit von 2% NMP. Die Enzymselektivität wurde dabei nicht negativ beeinflusst.

Table of contents

1 Abbreviations and acronyms

AA	Amino acid(s)
ACN	Acetonitrile
A. thaliana	*Arabidopsis thaliana*
BN-PAGE	Blue native polyacrylamide gel electrophoresis
BSA	Bovine serum albumin
B. gladioli	*Burkholderia gladioli*
DCM	Dichlormethane
DMF	N,N-Dimethylformamide
DMSO	Dimethylsulfoxide
DNA	Desoxyribonucleic acid
dATP	Deoxyadenosine triphosphate
dCTP	Deoxycytidine triphosphate
dGTP	Deoxyguanosine triphosphate
dTTP	Desoxythymidine triphosphate
dNTPs	Desoxynucleoside triphosphates
epPCR	error-prone polymerase chain reaction
ER	En(on)e reductase
E. coli	*Escherichia coli*
FMN	Flavin mononucleotide
GC	Gas chromatography
h	Hour(s)

HPLC	High performance liquid chromatography
kDa	Kilo Dalton
LB	Luria-Bertani / Lennox Broth
$MgCl_2$	Magnesium chloride
MES-NaOH	2-(N-Morpholino) ethanesulfonic acid sodium hydroxide
min	Minute(s)
mM	Millimolar
MTP	Microtiter plate
NAD(H)	β-Nicotinamide adenine dinucleotide (reduced disodium salt hydrate)
NADP(H)	β-Nicotinamide adenine dinucleotide 2-phosphate (reduced tetrasodium salt hydrate)
NMP	N-Methyl-2-pyrrolidone
PCR	Polymerase chain reaction
rpm	Rounds per minute
OD	Optical density
SDS-PAGE	Sodium dodecylsulfate polyacrylamide gel electrophoresis
SSM	Site saturation mutagenesis
S. virginiae	*Streptomyces virginiae*
Taq	*Thermus aquaticus*
TIM	Transposon Integration mediated Mutagenesis
TLC	Thin layer chromatography
U	Unit(s)
wt	Wild type

2 Introduction

New Methods and technologies for constructing diverse molecular libraries are a constant demand in order to generate and characterize new enzyme mutants. Several different strategies and tools to create functional diversity *in vitro* using recombination techniques or random mutagenesis and site-specific mutagenesis methods, respectively, are already well known (Arnold and Georgiou (Ed) 2003; Lutz and Patrick 2004) and routinely used in many laboratories. In nature insertion-deletion events additionally play an important role to create genetic diversity *in vivo*.

For this reason, in the first part of this study I report on a broadly applicable method for insertion/deletion mutagenesis as a tool for protein engineering. Transposon Integration mediated Mutagenesis (TIM) enables deletion or insertion of an arbitrary number of bases at random positions, insertion of functional sequence tags at random positions, replacing randomly selected triplets by a specific codon (e.g. scanning) and site-saturation mutagenesis. This transposon mutagenesis method combines random integration of modified bacteriophage *Mu* transposons with their subsequent defined excision employing type IIS restriction endonuclease *Aar*I. ENases-IIS have been implemented before in directed evolution experiments such as random insertion and deletion mutagenesis (RID) (Murakami *et al.* 2002), sequence-independent site-directed chimeragenesis (SISDC) (Hiraga and Arnold 2003) and the removal of triplet nucleotides at random positions (Jones 2005). The existing methods show some limitations e.g. RID is including a complex multistep procedure. As a proof of concept in this work a transposon named GeneOpener*Aar*IKan was designed and employed to introduce 6xHis tags randomly into the esterase EstC from *Burkholderia gladioli*. The employed strategy enables the isolation of randomly tagged active enzymes in single mutagenesis experiments. Some applications demonstrating the demand for internal tags were already published previously. For example Paramban and co-workers generated new GFP variants with an internal His tag for protein purification rationally. Those new GFP variants containing internal His tags can still be used for N-terminal and C-terminal fusion to proteins (Paramban *et al.* 2004). Randomly distributed introduction of functional amino acids such as cysteine, lysine or unnatural amino acids would enable to link enzymes onto carrier materials. This would further allow to identify the optimal orientation of an

immobilized enzyme e.g. to provide optimal access for the substrates. In this case all enzymes could be immobilized on carrier materials with the same optimized orientation. Immobilized enzymes have found numerous applications in analytical, clinical, environmental and industrial chemistry (Rao *et al.* 1998).

In the second part of this study I was pursuing the goal to apply protein engineering to generate improved enzymes from the industrially relevant enzyme class of non-heme oxidoreductases. As known, there is a demand for robust oxidoreductases with new or improved activities for the production of chiral intermediates for organic syntheses. Although they might be highly selective, heme dependent oxidoreductases often do not fulfil the criteria for successful applications under industrial conditions. Total turnover numbers are usually low, their coupling efficiency concerning use of costly redox coenzymes in relation to product formation is often low, and many of the known enzymes can only be employed for a limited range of substrates. Consequently, the aim in this part of the work was cloning, engineering and recombinant protein expression of new and/or interesting non-heme oxidoreductases in order to provide tailor-made biocatalysts for new industrial applications.

The first target for an enzyme engineering approach was the amine dehydrogenase from *Streptomyces virginae*. Itoh *et al*. claimed that this enzyme is capable of transforming various substrates. In particular, this enzyme was reported to be able to catalyze the reductive amination of ketones, not only keto acids (Itoh *et al.* 2000). As a consequence, such an amine dehydrogenase reveals a potential target for industry. Nevertheless, under industrial process conditions the wt enzyme still has some serious drawbacks e.g low stability or lack of enantioselectivity. Further engineering of the enzyme is therefore essential. Initially, the reproduction of the results from the paper of Itoh *et al.* was pursued. This included the establishment of different activity assays for indirect activity measurements and direct product detection.

The second enzyme of interest in this study was the 2-alkenal reductase P1 from *Arabidopsis thaliana.* This reductase is involved in preventing photooxidative stress in plants and specifically reduces the C=C bond of 2-alkenals. During DSM in-house screenings the enzyme from *A. thaliana* was active towards the non-natural α,β-unsaturated carbonyl compound 2-[4-methoxy-3-(3-methoxypropoxy) benzylidene]-3-methylbutanal. Due to the promising

results from the screenings, a codon optimized version of the gene was used as the starting point for enzyme engineering. Activity improvement of P1syn in the presence of organic co-solvents was focused in this work. P1syn was cloned into an appropriate *E. coli* strain for heterologous expression and designed to facilitate the industrial use. This part of the work included the establishment of a screening system and screening for new variants with enhanced activity towards 2-[4-methoxy-3-(3-methoxypropoxy) benzylidene]-3-methylbutanal in the presence of organic co-solvents, followed by purification of the wt enzyme as well as the best variant. Different mutagenesis strategies have been applied to improve the enzyme of interest.

Arnold, F. H. and G. Georgiou (Ed) (2003) *Directed Evolution Library Construction: Methods and Protocols* Humana Press, Totowa, New Yersey

Hiraga, K. and F. H. Arnold (2003) General Method for Sequence-independent Site-directed Chimeragenesis. *J. Mol. Biol.* **330**, 287-296.

Itoh, N., C. Yachi and T. Kudome (2000) Determining a novel NAD+-dependent amine dehydrogenase with a broad substrate range from Streptomyces virginiae IFO 12827: purification and characterization. *Journal of Molecular Catalysis B: Enzymatic*. **10**, 281-290.

Jones, D. D. (2005) Triplet nucleotide removal at random positions in a target gene: the tolerance of TEM-1 ß-lactamase to an amino acid deletion. *Nucleic Acids Res.* **33**, *e*80.

Lutz, S. and W. M. Patrick (2004) Novel methods for directed evolution of enzymes: quality, not quantity. *Current Opinion in Biotechnology*. **15**, 291-297.

Murakami, H., T. Hosaka and M. Sisido (2002) Random insertion and deletion of arbitrary number of bases for codon-based random mutation of DNAs. *Nat. Biotechnol.* **20**, 76-81.

Paramban, R. I., R. C. Bugos and W. W. Su (2004) Engineering green fluorescent protein as a dual functional tag. *Biotechnology and Bioengineering*. **86**, 687-697.

Rao, S. V., K. W. Anderson and L. G. Bachas (1998) Oriented Immobilization of Proteins. *Mikrochim. Acta*. **128**, 127-143.

2.1 Random tag insertions by Transposon Integration mediated Mutagenesis (TIM)

Brigitte M. Hoeller[1], Birgit Reiter[1], Sandra Abad[1], Ina Graze[1] and Anton Glieder*[1,2]

[1]Research Centre Applied Biocatalysis, Petersgasse 14, 8010 Graz, Austria

[2] Institute of Molecular Biotechnology, Graz University of Technology,

Petersgasse 14, 8010 Graz, Austria

* Corresponding author

Tel.: + 43 316 873 4074

Fax: + 43 316 873 4071

E-mail address: glieder@glieder.com

Keywords: ENases-IIS, EstC, GeneOpenerAarIKan, 6xHis tag, Mini-*Mu* transposon, Transposon mediated mutagenesis

2.1.1 Abstract

Transposon Integration mediated Mutagenesis (TIM) is a broadly applicable tool for protein engineering. This method combines random integration of modified bacteriophage *Mu* transposons with their subsequent defined excision employing type IIS restriction endonuclease *Aar*I. TIM enables deletion or insertion of an arbitrary number of bases at random positions, insertion of functional sequence tags at random positions, replacing randomly selected triplets by a specific codon (e.g. scanning) and site-saturation mutagenesis. As a proof of concept a transposon named GeneOpener*Aar*IKan was designed and employed to introduce 6xHis tags randomly into the esterase EstC from *Burkholderia gladioli*. A TIM library was screened with colony based assays for clones with an integrated 6xHis tag and for clones exhibiting esterase activity. The employed strategy enables the isolation of randomly tagged active enzymes in single mutagenesis experiments.

2.1.2 Introduction

Transposons are mobile genetic elements sometimes called "jumping genes" and were originally discovered in maize more than fifty years ago (McClintock 1987). Meanwhile they became indispensable tools in the genetic manipulation of prokaryotic as well as eukaryotic organisms (Haapa *et al.* 1999b). In nature, transposons move to different positions within the genome of a single host cell. They also cause many phenomena, such as spreading of antibiotic resistance among populations or the appearance of insertion and deletion mutations. Rearrangements mediated by transposons play a key role in remodelling chromosomes and thus are considered to be a driving force in genome evolution (Saedler and Gierl (Eds.) 1996). Apart from point mutations and recombination, nucleotide insertion and deletion is a common way to increase gene diversity in the natural process of evolution (Chothia *et al.* 2003). It was even suggested that nucleotide insertions and deletions, together with substitutions, represent the major mutational processes of gene evolution. Many insertion-deletion events in nature result in the addition or removal of single codons (Taylor *et al.* 2004). Such mutations mainly occur in coils and loops that link secondary structure elements (Pascarella and Argos 1992; Taylor *et al.* 2004).

In contrast, in typical directed evolution experiments the genetic diversity has been created so far almost exclusively by mutagenesis and/or recombination of one or more parent sequences (Neylon 2004). In 2002 a method termed random insertion/deletion (RID) mutagenesis was reported which enabled the deletion of an arbitrary number of consecutive bases at random positions and, at the same time, insertion of a specific sequence or random sequences of an arbitrary number into the same position (Murakami *et al.* 2002). However, since RID is a complex multistep procedure, our main goal was to employ a simple commercially available transposon and adapt it for insertion/deletion mutagenesis. A variety of transposon tools for sequencing and mutagenesis based on modified mini-*Mu* transposons are already commercially available. In laboratory, typical bacteriophage *Mu* applications include gene mapping, DNA sequencing, insertion mutagenesis and functional genetic analysis as well as functional analysis of proteins and protein-DNA complexes (Lamberg *et al.* 2002).

The bacteriophage *Mu* DNA transposition reaction mechanism is one of the most extensively studied (Haapa-Paananen *et al.* 2002) and bacteriophage *Mu* is the first bacteriophage for which an *in vitro* transposition reaction was shown (Mizuuchi 1983). An *in vitro Mu*

transposition reaction is relatively easy to perform and the only required macromolecular components are the *Mu*A transposase, the linear transposon DNA containing the two *Mu*A binding sites (R1 and R2), and the target DNA (Savilahti *et al.* 1995). To initiate transposition *in vitro*, four monomers of the *Mu*A transposase bind to the R1 and R2 recognition sites at the end of the transposon DNA to form the *Mu* transpososome. *Mu*A transposase makes 5 bp staggered cuts into the target DNA, and at the same time covalently joins the free 3´-OH of the transposon to the 5´-phosphate of the target site. Remaining single-stranded gaps are filled in by the *E. coli* DNA repair mechanism, resulting in the duplication of the 5 bp DNA target site (Allet 1979).

One of the most favourable characteristics for the use of bacteriophage *Mu* as a tool in molecular biology is its ability to insert with low target site selectivity, which produces relatively even distribution of transposon integrations into the target DNA (Haapa *et al.* 1999a; Haapa *et al.* 1999b).

TIM is based on a combination of random integration of modified bacteriophage *Mu*-mini transposons, the subsequent defined excision employing type IIS restriction endonuclease *Aar*I (Grigaite *et al.* 2002) and in the case of insertion mutagenesis the integration of oligonucleotides into the randomly opened DNA. The basic concept is very similar to the approach described by Jones 2005. They also engineered the mini-*Mu* transposon to serve as a carrier for the delivery of specific type IIS restriction endonuclease (ENase-IIS) recognition sequences into the target gene and thereby introduced a method for triplet nucleotide removal at random positions in a target gene to prove the tolerance of protein sequences to an amino acid removal (Jones 2005). ENases-IIS interact with two distinct sites on double stranded DNA: the recognition site, a 4-7 bp long asymmetric DNA sequence and the cleavage site, which is usually located between 1 and 20 bp from the recognition site (Roberts and Macelis 2001). Many unique applications of ENases-IIS have been described, including precise trimming of DNA, retrieval of cloned fragments, gene assembly, cleavage of single-stranded DNA, detection of point mutations, tandem amplification and localization of methylated bases (Szybalski *et al.* 1991). More recently ENases-IIS have been implemented in directed evolution experiments such as random insertion and deletion mutagenesis (RID) (Murakami *et al.* 2002), sequence-independent site-directed chimeragenesis (SISDC) (Hiraga and Arnold 2003) and as mentioned above the removal of triplet nucleotides at random positions (Jones 2005).

In this work, we designed a first broadly applicable mini-*Mu* transposon variant for random opening of plasmid DNA and for the consequent introduction of functional nucleotide sequences encoding a 6xHis tag. EstC (897 bp ORF) from *Burkholderia gladioli* was employed as a model protein to prove the flexibility of the method. Since this enzyme is well characterized, highly expressed in *E. coli* and mutant libraries can be easily screened with a high-throughput solid phase enzyme activity assay (Schlacher *et al.* 1998; Reiter *et al.* 2000).

2.1.3 Materials and Methods

2.1.3.1 *Bacterial strains, plasmids and culture conditions*

E. coli K12 TOP 10F´ F´(Youn *et al.* 2006) *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) ϕ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G (Invitrogen, Carlsbad, CA, USA) was used as a host for basic genetic work and library construction (TOP10 F´ Electrocomp™ Kits from Invitrogen). For esterase expression, the high copy number expression plasmid pMSEstC (Reiter *et al.* 2000) was used. This expression vector allows expression driven by the IPTG induced *tac* promoter.

Bacteria were grown at 37°C in Luria-Bertani (LB) broth or on LB agar plates (Sambrook *et al.* 1982). All components for *E. coli* media have been purchased from Carl Roth GmbH (Karlsruhe, Germany). For selection and maintenance of plasmids, antibiotics from Sigma Aldrich (St. Louis, MO, USA) were added to the growth media at the following final concentrations: 100 µg ml^{-1} ampicillin (Amp) and 10 µg ml^{-1} kanamycin (Kan). The mini-*Mu* Kan^R transposon was obtained from Invitrogen (GeneJumper™ Primer Insertion Kit for Sequencing.

2.1.3.2 *Enzymes and reagents*

Restriction endonucleases, Klenow fragment, T4 DNA Ligase, 50% PEG 4000 solution, Lambda DNA/*Hin*dIII Marker and GeneRuler™ 1kb DNA Ladder were from Fermentas (Burlington, ON, Canada). *Mu* transposase (*Mu*A) was obtained from Finnzymes (Espoo, Finland). *PfuUltra*™ High-Fidelity DNA Polymerase and Taq DNA Polymerase Buffer originated from Stratagene (La Jolla, CA, USA). Enzymes were used as recommended by the suppliers.

2.1.3.3 *Standard DNA techniques*

Kits from Promega GmbH (Madison, WI, USA) were used for DNA isolation from cells, agarose gels and for clean-up after enzymatic treatments. Nitrocellulose filters for desalting DNA were obtained from Millipore (Billerica, MA, USA). PCR reactions were performed in a GeneAmp® PCR System 2400 from Applied Biosystems (Foster City, CA, USA). Direct PCR Cloning was done with Zero Blunt® TOPO® PCR Cloning Kit (Invitrogen).

2.1.3.4 *Construction of pCRGeneOpenerAarIKan*

A scheme for the construction of pCRGeneOpener*Aar*IKan is outlined in Figure 2.1. In a first step, the *Aar*I sites were integrated into the GeneJumper™ KanR Transposon using single primer PCRs with the primers *Mu*ins1bp*Aar*I (5'-GCTCAGATCTGGCAGGTGGCACGAAAAACG-3') or *Mu*del1bp*Aar*I (5'-GGAGCTCAGATCTGCAGGTGCGCACGAAAAACG-3'). This resulted in the transposons ins1bp*Aar*I and del1bp*Aar*I, respectively. In addition to the *Aar*I recognition site these primers contained a *Bgl*II site, for the generation of a linear transposon DNA. For the PCRs an initial denaturation step of 5 min at 95°C was employed followed by 25 cycles of 30 sec at 94°C, 30 sec annealing at 52°C and 2 min extension at 72°C and a final elongation step of 10 min at 72°C. PCR reaction mixtures (50 µl) contained 0.4 µM primer, 1 x PfuUltra High-Fidelity reaction buffer, 200 µM each of dATP, dCTP, dGTP and dTTP, 2 U PfuUltra® High-Fidelity DNA Polymerase and 20 ng GeneJumper™ KanR transposon as template. PCR products were purified via agarose gel electrophoresis followed by gel elution with the Wizard® SV Gel and PCR Clean-Up System and then directly amplified in *E. coli* using the pCR-Blunt II TOPO cloning vector. This resulted in the plasmids pCRins1bpAarIKan and pCRdel1bp*Aar*IKan. The final gene opener plasmid pCRGeneOpener*Aar*IKan was gained by combining the 5´-end of the transposon del1bp*Aar*I and the 3´-end of ins1bpAarI. Therefore, both plasmids pCRins1bp*Aar*IKan and pCRdel1bpAarIKan were digested with *Not*I, which cuts twice, inside the transposon as well as downstream of the transposon DNA in the multiple cloning site of the vector. Using 3 U of T4 DNA ligase the small *Not*I fragment of pCRins1bpAarIKan was ligated into the residual vector backbone of pCRdel1bp*Aar*IKan. After transformation the sequence of the final construct pCRGeneOpener*Aar*IKan was verified by DNA sequencing.

2.1.3.5 *Transposition and creation of TIM libraries with randomly integrated His tags*

A general overview of TIM is shown in Figure 2.2. pCRGeneOpener*Aar*IKan was digested with *Bgl*II, the 1.2 kb transposon band was gel purified and employed for *in vitro* transposition. 200 ng target plasmid pMSEstC were mixed with 4 µl 5x reaction buffer, 20 ng transposon GeneOpener*Aar*IKan, 1 µl *Mu*A Transposase (0.22 µg μl^{-1}) and ddH_2O in a final volume of 20 µl. The reaction mixture was incubated for 2 h at 30°C. 1 µl of the transposition mixture was used for transformation by electroporation. Successful insertion of GeneOpener*Aar*IKan into pMSEstC conferred resistance of *E. coli* to kanamycin. Thus, plasmids with an integrated transposon were selected by growth of the host cells on media containing both ampicillin and kanamycin. 5 µg transposed plasmid DNA, 5 µl 10x reaction buffer, 1 µl 0.025 mM oligonucleotide containing the *Aar*I recognition sequence, 2 µl *Aar*I (3 U μl^{-1}), and ddH_2O in a final volume of 50 µl were digested to release the integrated gene opener transposons (GeneOpener*Aar*IKan).

The linearized plasmid pMSEstC (~ 5.2 kb) was gel purified and employing Klenow polymerase blunt ends were generated at the 5´ single-strand extensions. Therefore a reaction containing ~ 3 µg purified plasmid DNA was performed to fill-in recessed 3’ termini of double stranded DNA. The DNA was purified again using the Wizard® SV Gel and PCR Clean-Up System and the final step comprised the ligation of oligonucleotides encoding 6xHis tags into the blunt-end opened plasmids employing T4 DNA ligase in the presence of 5% polyethylene glycol. Therefore 30 µl blunt-ended linear plasmid DNA were incubated over night at 16°C with the linkers 6xHisG (5’- GGCCATCATCATCATCATCACGGC-3’) and 6xHisGrev (5’-GCCGTGATGATGATGATGATGGCC-3’) in a concentration of 3 µM using 1 µl T4 DNA ligase (5 U µl-1), 5 µl 50% PEG4000 solution, 5 µl 10x ligation buffer and ddH2O in a final volume of 50 µl. After ligation the T4 DNA ligase was inactivated, the ligated DNA was purified and eluted from the purification column with 180 µl nuclease free water. Subsequently, 20 µl 10x Taq Polymerase Buffer were added, in order to stabilize the DNA during the following heating step. The DNA was heated at 80°C for 10 min, cooled down to room temperature, purified again and eluted from the purification column with 30 µl nuclease free water. 3 µl of this ligated DNA sample was transformed.

2.1.3.6 *Colony-blot screening*

The colonies of transformants growing on LB Amp agar plates were blotted onto nitrocellulose membranes. EstC expression was induced by transferring the membranes onto LB plates containing ampicillin and 500 µM IPTG. Before 6xHis tag-containing proteins were detected using QIAexpress Anti·His HRP Conjugates, colonies were grown on these plates for about 4 h. After this the membranes (colony side up) were incubated with lysis buffer containing 0.1 M potassium phosphate buffer (pH 8.0), 1 mg ml^{-1} lysozyme, 0.5 µg ml^{-1} DNaseI, 10 mM $MgCl_2$ and 0.1% Tween 20. The colonies were allowed to lyse for 1h at 37°C. His tag screening was performed according to the QIAexpress® Detection and Assay Handbook with two exceptions. For the washing step TBS Tween Buffer was used and the Anti·His HRP Conjugate blocking buffer contained 0.05% (v/v) Tween 20 (final concentration). A Hybaid mini hybridization oven was employed for the immunogenic reactions at room temperature (Thermo Electron Corporation, Waltham, MA, USA). Anti·His HRP Conjugates allowed direct detection of 6xHis-tagged enzymes by a chromogenic method using 4-chloro-1-naphthol (4C1N) as a substrate. Consequently, clones carrying a 6xHis tag were analyzed for esterase activity. Colony-blot screening was performed once. Approximately 4400 colonies were screened to exemplify a proof of concept of Transposon Integration mediated Mutagenesis.

2.1.3.7 *Esterase screening*

Esterase producing colonies were directly identified on plates with a simple filter assay as described by Schlacher (Schlacher *et al.* 1998). Therefore the filters (Filter Paper Circles MN 619 eh 9 cm diameter) (Macherey-Nagel, Düren, Germany) were soaked with a substrate/indicator solution (3 mM alpha-naphthyl acetate, 1 mM fast blue B (Sigma), 100 mM Tris-HCl buffer, pH 7.0) and subsequently directly placed onto the colonies grown on LB Amp agar plates. Colonies expressing active EstC developed a deep-purple color within a few seconds.

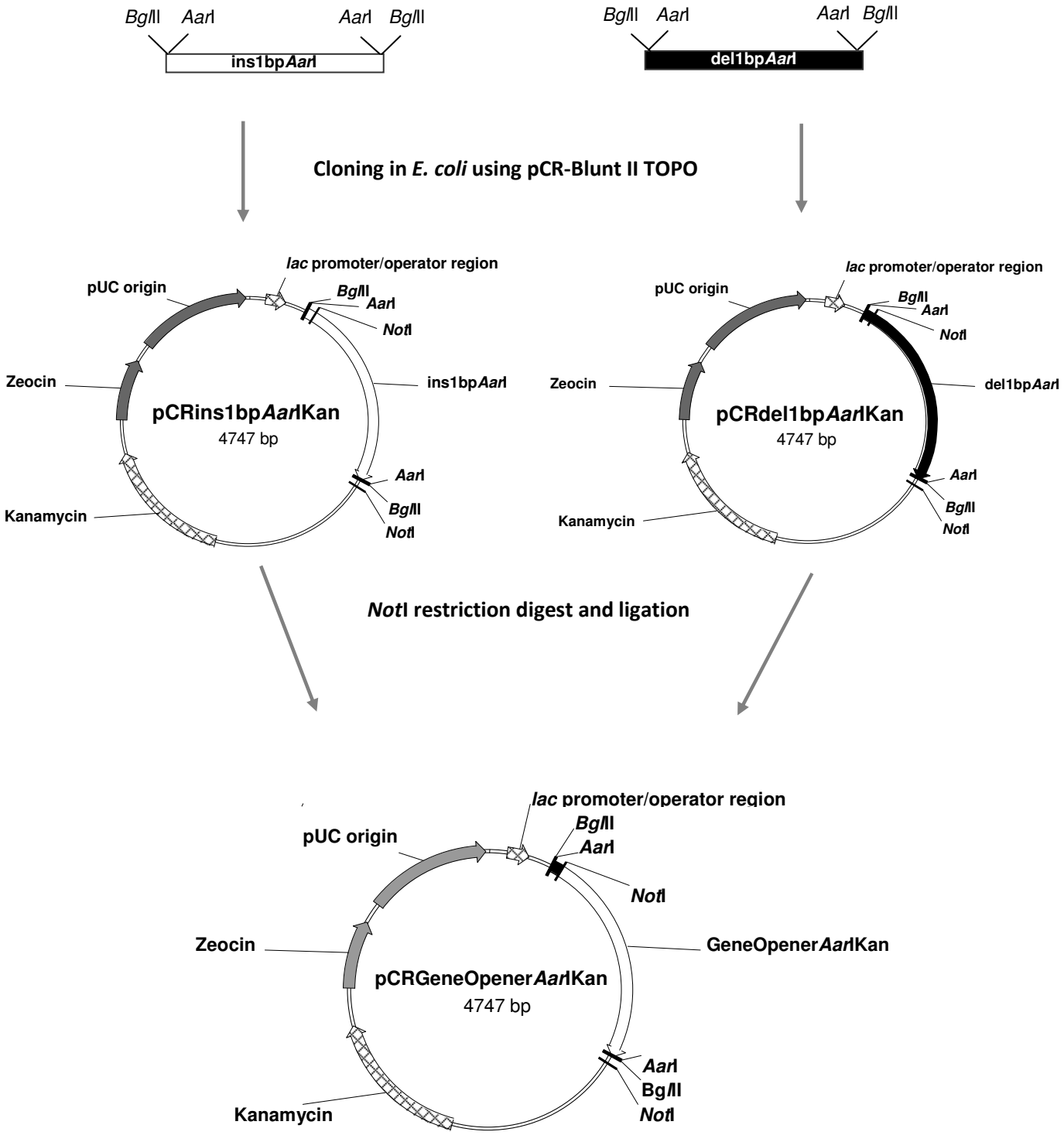
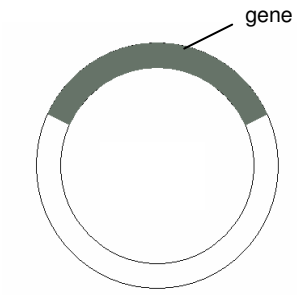
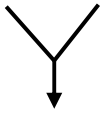


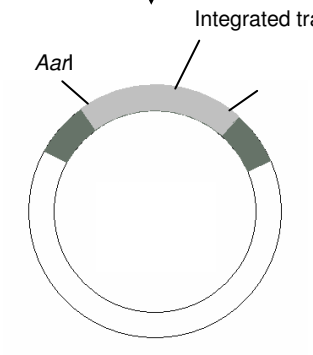
Figure 2.1: Schematic representation of the construction of the engineered gene opener transposon for TIM. The transposons ins1bp*Aar*IKan and del1bpAarIKan containing the *Aar*I recognition sites were directly amplified in E. coli using the vector pCR-Blunt II TOPO. The resulting plasmids pCRins1bp*Aar*IKan and pCRdel1b*Aar*IKan were digested with *Not*I. Finally, transposon GeneOpener*Aar*IKan was created by combining the small *Not*I fragment of pCRins1bp*Aar*IKan with the residual vector backbone of pCRdel1bp*Aar*IKan.



-TRANSPOSON INTEGRATION MEDIATED MUTAGENESIS-

Transposition





Selection of kanamycin resistant clones

Nucleotide sequence level

1 5´–XXXXXX**12345**T*GCAGGTG*CGCACG --- CGTGC*CACCTGC*CA**12345**XXXXXX–3´

*Aar*I digest

5´–XXXX **345**XXXXXX–3´

3´–XXXXXX**12** XXXXX–5´

2

Klenow fill-up  6xHis linker ligation

3 5´–XXXXXX**12**6xHis**345**XXXXXX–3´

3´–XXXXXX**12**6xHis**345**XXXXXX–5´

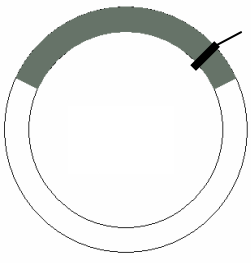


Figure 2.2 Experimental outline of Transposon Integration mediated Mutagenesis (TIM). The first of 4 steps is the transposition reaction for random integration into the target DNA. The transposons transferring resistance towards kanamycin are randomly integrated into plasmid DNA carrying the ampicillin resistance gene. The transposition reaction mixture is directly transformed into competent *E. coli* cells. Plasmids containing a transposon are selected by growth on LB plates containing ampicillin and kanamycin. Transformants are carefully scratched off the plates using a spatula and total plasmid DNA is isolated. Since the transposon encodes *Aar*I recognition sites on each end, *Aar*I digest releases the transposon out of the plasmid again. Nucleotide sequence level: In lane 1 the site of integration of the transposon with the 5 bp duplications (bold) and the *Aar*I recognition sites (italics) are shown. The Xs indicate the flanking gene sequences. The transposon is released from the target DNA by restriction with the type IIS restriction endonuclease *Aar*I. *Aar*I cuts 4 bp and 8 bp away from its recognition site leaving 4 bp long 5´-protruding ends shown in lane 2. Lane 3: The overhangs are filled up to blunt ends using Klenow fragment. Finally, the 6xHis tag coding sequences are ligated into the linearized, blunt-ended plasmids and transformed into competent *E. coli* cells. Transformants are selected by ampicillin resistance and screened for esterase activity and for randomly integrated 6xHis tags.

2.1.4 Results

2.1.4.1 *Transposition reaction and target site selection*

Transformation of 1 µl of the transposition reaction mixture into *E. coli* cells with a competence of 10^8-10^9 cfu μg^{-1} DNA resulted in typically about 10^4-10^5 cfu per reaction. For a rough analysis of target site preference of the mini-*Mu* transposon variant GeneOpener*Aar*IKan 10 colonies carrying transposed plasmids were selected and transposon integration sites were determined by DNA sequencing with Primer B of the GeneJumper™ Primer Insertion Kit for Sequencing. The obtained sequences were aligned with the wild-type sequence of pMSEstC in order to visualize the distribution of the GeneOpener*Aar*IKan insertions. GeneOpener*Aar*IKan transposons utilized in this study were distributed throughout the entire plasmid, except within the essential origin. Expectedly, also no transposons were inserted into the Amp resistance cassette, because insertion of the transposon would lead to inactivation of the selection marker.

2.1.4.2 *Colony selection and TIM based sequence analysis*

To prove the concept of random integration of 6xHis tags into EstC from B. gladioli, a TIM library was screened to find a functional 6xHis-tagged clone using Anti·His HRP Conjugates in a colony-blot procedure. Approximately 4400 colonies were screened to detect a 6xHis tag positive clone by developing a blue color due to an insoluble, colored product that formed a precipitate on the membrane. The clone also exhibited esterase activity as determined by the simple filter assay mentioned above. This clone named TIM1 was sequenced using the primers TacNEU (5´-GAATTCGAGCTCGGTAC-3´) and TacSTOP (5´-CATCCGCCAAAACAGCC-3´) which bind in the 5´ and 3´ flanking regions of *estC*. Thus, the whole esterase C gene was sequenced

and thereby the exact position of the incorporated 6xHis tag coding sequence in the target gene was determined. Sequence analysis of the 6xHis tag positive clone TIM1 revealed that the 6xHis tag was inserted into the target gene at the outermost 3´ end of *estC*. In TIM1, the *est*C stop codon was shifted by a 1 bp deletion in the target gene, thereby prolonging the translated polypeptide up to the next in-frame stop codon downstream (39 bases distal of the original one). Apparently, this did not disturb protein folding or esterase activity (Figure 2.3a).

In addition to His tag screening some 1500 colonies of the library were investigated for esterase activity using the simple filter paper solid phase assay. Approximately 1200 colonies showed esterase activity, since the transposon can integrate randomly into the whole plasmid. 6xHis tag coding oligonucleotides were not exclusively integrated into the coding DNA. Naturally, integration into both the ori or selection marker prevents growth. A rough estimation of the distribution of 6xHis tags throughout the target plasmid pMSEstC was done. For this reason, 28 esterase-positive clones and for comparison 5 clones showing no esterase activity were analyzed by sequencing with the primer histagTIMfw (5´-GCCATCATCATCATCATCAC-3´) (Figure 2.4). For 3 out of the 28 sequenced clones showing esterase activity no sequence was observed. Obviously, these plasmids did not integrate 6xHis linkers during blunt-end ligation into the opened plasmid. From the remaining 25 esterase active clones another sequence (named TIM2) was found harboring the 6xHis tag coding sequence within the esterase gene. For determination of the exact position of the incorporated 6xHis tag coding sequence in the target gene the primers TacNEU and TacSTOP were used again for sequencing. The 6xHis tag coding oligonucleotide was integrated at the 3' end of the esterase gene but upstream of the stop codon. In addition it was reverse oriented, resulting in Arg and 5xAsp instead of 6xHis, in this particular case. Thus, this clone could not be detected in the colony-blot procedure for 6xHis tag detection. Interestingly, although a deletion at the very end of the tag or a deletion in the plasmid (C is missing) caused a frame shift, esterase activity of this interrupted ORF was still retained. Translation most probably gets terminated by a new in-frame stop codon further downstream (42 bases distal of the original one) (Figure 2.3b). All 5 clones (TIM-1, TIM-2, TIM-3, TIM-4 and TIM-5) lacking esterase activity harbored a 6x His coding sequence in the target gene *estC*, but at different sites of the gene. In all 5 cases tag incorporation led to a frame shift, partly by deletion of one base in the sequence straight in front of the 6xHis tag integration site (Figure 2.3c). In the

cases of TIM-1, TIM-2 and TIM-3 incorporation of the designed 6xHis tag occurred in reverse orientation.

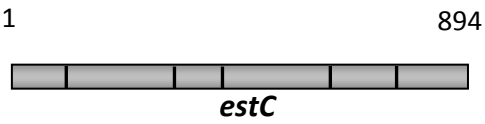
a

```
pMSEstC          CACGGCGATC----------------------TGAGGCAGCGATCGAGCGGCCCGTGT
TIM1             CACGGCGATCGGCCATCATCATCATCATCACGGCGAGGCAGCGATCGAGCGGCCCGTGT
                 **********                      ************************

pMSEstC          CCGCGAGTCGTTTATGATGACGATCATTCGATGACAGGCCGCCGCACGCTGCAGCTCCC
TIM1             CCGCGAGTCGTTTATGATGACGATCATTCGATGACAGGCCGCCGCACGCTGCAGCTCCC
                 ************************************************************
```

b

```
pMSEstC          TGACGGGCATCGCGCGCAACA----------------------CGGCGATCTGAGGCAG
TIM2             TGACGGGCATCGCGCGCAACAGCCGTGATGATGATGATGGCCGGCGATCTGAGGCAG
                 **********************                     ***************

pMSEstC          CGATCGAGCGGCCCGTGTCCGCGAGTCGTTTATGATGACGATCATTCGATGACAGGCCGC
TIM2             CGATCGAGCGGCCCGTGTCCGCGAGTCGTTTATGATGACGATCATTCGATGACAGGCCGC
                 ************************************************************
```

c



Clone	Insert site (at position)	Insertion	Frame shift
TIM-1	135bp	GCCGTGATGATGATGATGATGGC	+
TIM-2	360bp	GCCGTGATGATGATGATGATGGC	+
TIM-3	427bp	GCCGTGATGATGATGATGATGGCC	+ / 1bp deleted
TIM-4	701bp	GGCCATCATCATCATCATCACGGC	+
TIM-5	597bp	GGCCATCATCATCATCATCACGGC	+ / 1bp deleted

Figure 2.3 Alignment of pMSEstC wild-type sequence and the sequences of the two TIM clones showing esterase activity and harboring a 6xHis tag within the *estC* gene. Stars indicate sequence identity. Stop codons are framed. (a) 6xHis linker incorporation at the C-terminal end of the protein EstC. The integration of the 6x His tag and a deletion of 1bp in the target gene led to a displacement of the stop codon. (b) 6xHis linker incorporation at the C-terminal end of the protein EstC. In this case the 6xHis coding sequence integrated in reverse orientation. Potentially, a deletion of 1 bp at the very end of the tag or in the plasmid (a C is missing) caused a frame shift and a displacement of the stop codon, while retaining esterase activity (c) Integration sites and 6xHis coding sequences of the 5 TIM clones showing no esterase activity. + out of frame. Bars indicate the insertion sites.

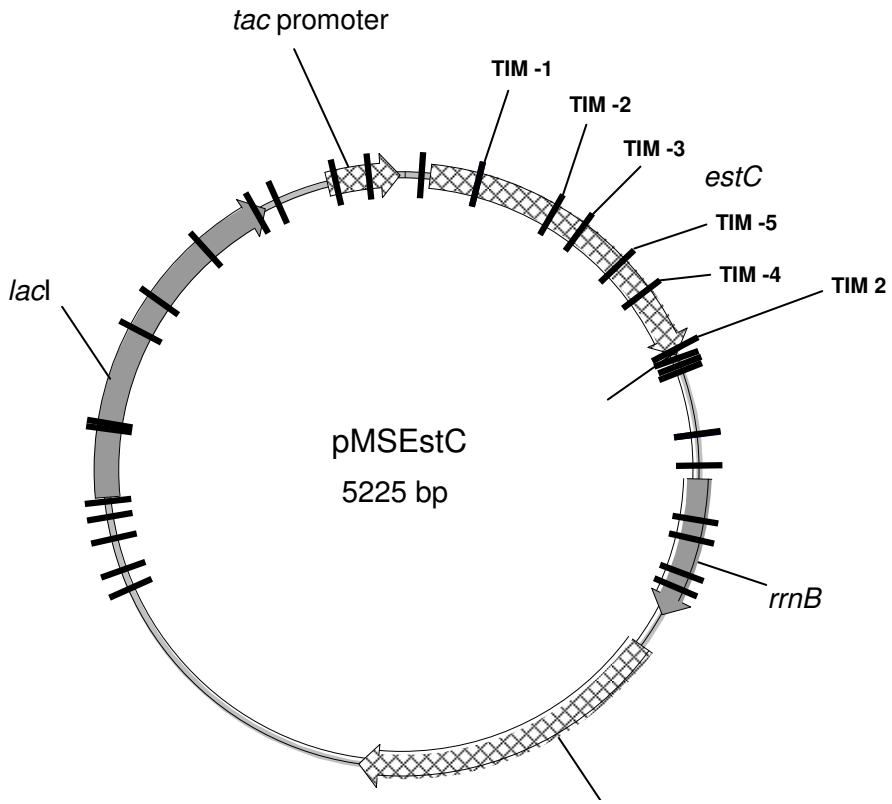


Figure 2.4 Distribution of the 6xHis tag coding sequences throughout the target plasmid pMSEstC determined by sequencing. The bars indicate the 6xHis linker integration sites. TIM clones harboring a 6xHis linker sequence in the target gene *estC* are labelled. The *estC* gene, the *lacI* gene, the ampicillin resistance gene, the *tac* promoter and the *rrnB* site are indicated as arrows.

2.1.5 Discussion

Recently, Jones described a directed evolution method for the introduction of triplet nucleotide deletions at random positions throughout a target gene (Jones 2005). The author employed engineered mini-*Mu* transposons to introduce recognition sequences for the ENase-IIS *Mly*I. However, this particular enzyme generates blunt ends by cutting 5 bp outside of the recognition sequence GAGTC which occurs frequently in any DNA sequence, thus limiting simple applications of insertion/deletion mutagenesis in enzyme breeding projects. This shortcoming for an implementation of this technique in a directed evolution program was overcome by using longer and less common recognition sequences. Consequently, also larger genes and whole plasmids can be mutated by TIM, without prior removal of statistically frequent 5 bp restriction enzyme recognition sites.

Unfortunately, *Mly*I currently is the only commercially available ENase-IIS that cuts double stranded DNA in a blunt-end manner. In our study we were pursuing the strategy to apply a restriction enzyme which generates sticky ends but uses a longer and thus statistically less common recognition sequence. ENase-IIS *Aar*I cuts in an appropriate distance from the recognition site and generates 5´-protruding ends (Grigaite *et al.* 2002). Blunt ends are then obtained by a fill-up reaction with DNA polymerases (e.g. Klenow fragment, T4 polymerase).

As a proof of concept pMSEstC harboring the esterase gene from *B. gladioli* acted as target DNA for the insertion of GeneOpener*Aar*IKan transposon containing *Aar*I recognition sites at both ends of the transposon. For *Aar*I it was known that efficient restriction of DNA requires the presence of two or more recognition sites in an optimal relative distance (Bath *et al.* 2002; Grigaite *et al.* 2002). The addition of 0.5 µM oligonucleotide containing the *Aar*I recognition sequence increased the yield of cut DNA and a further purification step by gel electrophoresis made the presented method more efficient.

In a few cases we observed deletions of exactly 1 base directly in front or right after the 6xHis coding oligonucleotide insertion sites. According to literature type IIS restriction endonucleases cleave both DNA strands at fixed positions (Bath *et al.* 2002), although we have no proof, we speculate, *Aar*I is not always cutting exactly, which is in contrast to our application including high-throughput screening not detected during standard ligation procedures.

Nevertheless, the engineered gene opener transposon provides a generally applicable and simple tool and thus the basis for the introduction of any additional sequence such as e.g. 6xHis tags within a target DNA. This provides an alternative to targeting each possible integration site individually by site-directed mutagenesis. Employing a single random based mutagenesis experiment combined with screening, this method serves as an efficient tool to identify optimal positions in biomolecules where functional groups, tags or even stop codons can be introduced. Some applications demonstrating the demand for internal tags were published previously. For example Paramban *et al.* (2004) generated new GFP variants with an internal His tag for protein purification rationally. Those new GFP variants containing internal His tags can still be used for N-terminal and C-terminal fusion to proteins (Paramban *et al.* 2004). TIM can also be used for randomly distributed introduction of functional amino acids such as cysteine or unnatural amino acids which enable to link enzymes onto carrier materials.

This would allow to identify the optimal orientation of an immobilized enzyme e.g. to provide optimal access for the substrates. In this case all enzymes could be immobilized on carrier materials with the same optimized orientation. Immobilized enzymes have found numerous applications in analytical, clinical, environmental and industrial chemistry (Rao *et al.* 1998), but typically there is a significant loss of enzymatic activity due to immobilization. Furthermore, random integration of N-glycosylation sites (N x T/S) into secreted proteins expressed by eukaryotic hosts would be another interesting application which can be evaluated using TIM and would enable to alter surface properties, solubility and stability of enzymes.

Summarizing, we proved that TIM is a simple method for the introduction of tags into proteins and due to the rare 7 bp recognition site of AarI it is also broadly applicable. Placing additional oligonucleotides into our model protein EstC employing a new artificially designed mini-transposon appeared randomly and with high efficiency and the 6xHis-tag-encoding 24 bp-cassette was distributed throughout the target plasmid. Screening only 4400 colonies using a simple solid phase assay we were able to detect clones containing 6xHis tag coding sequences in the target protein EstC. One of the clones expressing EstC with an incorporated 6xHis tag in the right frame was also active. A further variant at least indicated another possible position of integration into the model enzyme EstC, without subsequent loss of enzymatic activity. TIM will be useful for many further applications like serine, alanine or cysteine scanning, domain shuffling to create chimeric proteins and for random truncation of proteins by introduction of tags combined with stop codons.

2.1.6 Acknowledgements

We thank the FFG, DSM, the province of Styria as well as SFG for financial support. Special thanks to Roland Weis for his advice and critical reading of the manuscript.

2.1.7 References

Allet, B., 1979. Mu insertion duplicates a 5 base pair sequence at the host inserted site. Cell 16, 123-129

Bath, A. J., Milsom, S. E., Gormley, N. A., Halford, S. E., 2002. Many Type IIs Restriction Endonucleases Interact with Two Recognition Sites before Cleaving DNA. J. Biol. Chem. 277, 4024-4033

Chothia, C., Gough, J., Vogel, C., Teichmann, S. A., 2003. Evolution of a protein repertoire. Science 300, 1701-1703

Grigaite, R., Maneliene, Z., Janulaitis, A., 2002. AarI, a restriction endonuclease from *Arthrobacter aurescens* SS2-322, which recognizes the novel non-palindromic sequence 5'-CACCTGC(N)4/8-3'. Nucl. Acids Res. 30, *e*123

Haapa, S., Suomalainen, S., Eerikäinen, S., Airaksinen, M., Paulin, L., Savilahti, H., 1999a. An Efficient DNA Sequencing Strategy Based on the Bacteriophage Mu in Vitro DNA Transposition Reaction. Genome Res. 9, 308-315

Haapa, S., Taira, S., Heikkinen, E., Savilahti, H., 1999b. An efficient and accurate integration of mini-Mu transposons *in vitro*: a general methodology for functional genetic analysis and molecular biology applications. Nucleic Acids Res. 27, 2777-2784

Haapa-Paananen, S., Rita, H., Savilahti, H., 2002. DNA transposition of bacteriophage Mu. A quantitative analysis of target site selection *in vitro*. J. Biol. Chem. 277, 2843-2851

Hiraga, K., Arnold, F. H., 2003. General Method for Sequence-independent Site-directed Chimeragenesis. J. Mol. Biol. 330, 287-296

Jones, D. D., 2005. Triplet nucleotide removal at random positions in a target gene: the tolerance of TEM-1 ß-lactamase to an amino acid deletion. Nucleic Acids Res. 33, *e*80

Lamberg, A., Nieminen, S., Qiao, M., Savilahti, H., 2002. Efficient Insertion Mutagensis Strategy for Bacterial Genomes Involving Electroporation of In Vitro-Assembled DNA Transposition Complexes of Bacteriophage Mu. Appl. Environ. Microbiol. 68, 705-712

McClintock, B., 1987. The discovery and characterization of transposable elements: the collected papers of Barbara McClintock. Genes, Cells, and Organisms. Garland Pub., New York, NY

Mizuuchi, K., 1983. In *vitro* transposition of bacteriophage Mu: A biochemical approach to a novel replication reaction. Cell 35, 785-794

Murakami, H., Hohsaka, T., Sisido, M., 2002. Random insertion and deletion of arbitrary number of bases for codon-based random mutation of DNAs. Nat. Biotechnol. 20, 76-81

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

Paramban, R. I., Bugos, R. C., Su, W. W., 2004. Engineering green fluorescent protein as a dual functional tag. Biotechnology and Bioengineering 86, 687-697

Pascarella, S., Argos, P., 1992. Analysis of insertions/deletions in protein structures. J. Mol. Biol. 224, 461-471

Rao, S. V., Anderson, K. W., Bachas, L. G., 1998. Oriented Immobilization of Proteins. Mikrochim. Acta 128, 127-143

Reiter, B., Glieder, A., Talker, D., Schwab, H., 2000. Cloning and characterization of EstC from *Burkholderia gladioli*, a novel-type esterase related to plant enzymes. Appl. Microbiol. Biotechnol. 54, 778-785

Roberts, R. J., Macelis, D., 2001. REBASE-restriction enzymes and methylases. Nucleic Acids Res. 29, 268-269

Saedler, H., Gierl, A., (Eds.) 1996. Transposable elements. Springer-Verlag, Berlin, Heidelberg.

Sambrook, J., Fritch, E. F., Maniatis, T., 1982. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y.,

Savilahti, H., Rice, P. A., Mizuuchi, K., 1995. The phage Mu transpososome core: DNA requirements for assembly and function. EMBO J. 14, 4893-4903

Schlacher, A., Stanzer, T., Osprian, I., Mischitz, M., Klingsbichel, E., Faber, K., Schwab, H., 1998. Detection of a new enzyme for stereoselective hydrolysis of linalyl acetate using simple plate assays for the characterization of cloned esterases from Burkholderia gladioli. Journal of Biotechnology 62, 47-54

Szybalski, W., Kim, S. C., Hasan, N., Podhajska, A. J., 1991. Class-IIS restriction enzymes - a review. Gene 100, 13-26

Taylor, M. S., Ponting, C. P., Copley, R. R., 2004. Occurence and consequences of coding sequence insertions and deletions in mammalian genomes. Genome Res. 14, 555-566

2.2 Enzymatic reductive amination of ketones

2.2.1 State of the art

Amino acid dehydrogenases converting keto acids to the corresponding amino acids employing ammonia and vice versa have been described many times and are well known. There are several publications about amino acid dehydrogenases showing reductive amination activities (e.g. L-leucin dehydrogenase, L-phenylalanin dehydrogenase). 3D structures of such enzymes are available (Hummel and Kula 1989; Baker *et al.* 1995). Some of them have already been implemented on commercial scale, e.g. by Degussa for the preparation of L-*tert*-leucine. In general, naturally available amino acid dehdrogenases exclusively display activity with alpha keto acids, not ketones.

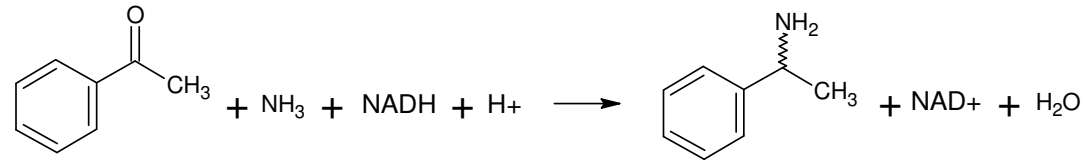
DSM brought in the information about an interesting amine dehydrogenase, which according to Itoh *et al.* allows the production of amines from ketones. In addition, an engineered dehydrogenase which allows reductive amination of ketones was announced on the homepage of the German company X-Zyme. One publication and two patent applications (US2003027306 A1, US6432688 B1) have been disclosed by the Japanese group (Itoh *et al.* 2000; Ito *et al.* 2002-08-13; Ito *et al.* 2003-02-06) which reported the discovery and purification of the novel amine dehydrogenase (AMDH) from *Streptomyces virginiae* IFO 12827. According to the publication this AMDH has an extremely broad substrate specificity and can catalyze the reductive amination (NADH-dependent) of keto acids, keto alcohols, ketones and aldehydes. However, the enzyme has also a number of serious defects, e.g. its low stability and (more important) lack of enantioselectivity. This makes the enzyme unsuitable for the production of chiral amines and amino alcohols for commercial scale production. Besides that, during the last 9 years there was not any follow-up study described by Itoh and co-workers, causing some doubts about the reproducibility of their very interesting findings. In addition, the nucleotide/amino acid sequence has not been disclosed.

2.2.2 Objectives

The reaction of interest in our case was the enzymatic reductive amination of ketones. In particular, we wanted to investigate the activity of the amine dehydrogenase towards the model substrates acetophenone and 2-octanone. Therefore, my main goal in the first place was to reproduce the findings claimed in the paper and patents from Itoh and co-workers.

Model reactions for enzymatic reductive aminations with the AMDH from *Streptomyces virginiae*:

a) Aromatic – acetophenone



b) Aliphatic – 2-octanone

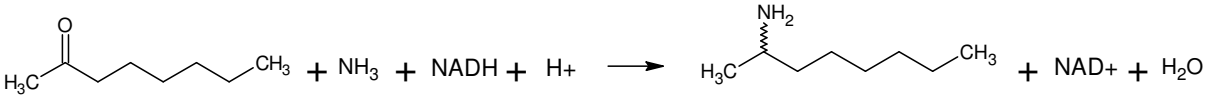


Figure 2.5 Reductive amination of a) 2-octanone to 2-octylamine and b) acetophenone to 1-phenylethylamine

2.2.3 Materials and methods

2.2.3.1 *Strain, culture conditions and cell disruption*

The *Streptomyces virginiae* IFO 12827 strain (DSMZ, Germany) was used throughout this study. The media for pre-culture (50 ml, 2 days at 28°C, 110 rpm) and main-culture (100 ml, 36 h at 28°C, 110 rpm) were made according to Itoh and co-authors (Itoh *et al.* 2000). The recipe of the TM solution was gathered from the United States patent US6432688 B1. The cells were harvested from 100 ml cultures by centrifugation at 4°C for 15 min at 4,000 rpm (Beckman Coulter JA 10) and stored at -20°C for further use. The cell pellets were resuspended in 5 mL of the respective reaction buffer. Subsequently, the cells were lysed via sonication (Branson Sonifier 250, Inula, Vienna) for 14 min under constant cooling. Finally, the lysates were transferred to Eppendorf tubes and solid cell debris were removed by centrifugation at 4°C for 30 min at 13,200 rpm. The supernatant was further examined. As a

control strain *E. coli* BL21-AI (Invitrogen) was cultivated in LB broth containing 10 g L^{-1} tryptone, 5 g L^{-1} yeast extract and 5 g L^{-1} NaCl (same culture volumes as used for the *S. virginae*). Cell disruption was performed as already described.

2.2.3.2 *Chemicals and reagents*

Serinol (2-amino-1,3-propanediol), INT, ninhydrin, acetophenone, 2-octanone, (*R*)-1-phenylethylamine, *rac*-1-phenylethylamine and 2-octylamine were purchased from Sigma Aldrich (St. Louis, MO, USA). Potassium permanganate and ethanol were from Carl Roth (Karlsruhe, Germany).

2.2.3.3 *Determination of protein concentrations*

Total protein concentrations were determined by BCA Protein Assay Kit (Novagen, United States).

2.2.3.4 *Photometric enzyme assays*

AMDH Activity – Oxidative deamination reaction

Oxidative deamination was spectrophotometrically assayed at 25°C by measuring the increase in absorption of formazan produced from 2-(4-iodophenyl)-3-(4-nitrophenyl)-5-phenyl-2H tetrazolium chloride (INT) at 490 nm. The reduction of NAD^+ to NADH leads to the reduction of INT to a highly-colored red-purple formazan. The reaction mixture consisted of 0.1 M Tris HCl buffer (pH 8.0) containing 0.5 mg mL^{-1} INT, 10 mM of the respective substrate, 1 mM NAD^+ and 50 µL of cell free extract in a total volume of 1.5 ml. The reaction was started by adding the cell free extract. The blank reaction contained buffer instead of substrate.

AMDH Activity – Reductive amination reaction

Reductive amination activity of the AMDH was measured by NADH decrease at 340 nm. The reaction mixture consisted of 0.1 M potassium phosphate buffer (pH 7.0), 10 mM dihydroxyacetone dimer, 0.2 M NH_4Cl, 0.1 mM NADH and 50 µL of lysates (*S. virginiae*, *E. coli* BL21-AI) in a total volume of 1.5 ml. As mentioned before, the blank contained buffer instead of substrate.

2.2.3.5 *Analytical methods for model reactions*

TLC

For the $KMnO_4$ staining solution 0.5% (w/v) potassium permanganate was dissolved in water and the ninhydrin staining solution contained 0.3% (w/v) ninhydrin in ethanol. For aromatic substrates detection was carried out under UV light and for aliphatic substrates detection was performed with ninhydrin and $KMnO_4$. Reaction mixtures for oxidative deamination contained ~30 mM of each substrate ((*R*)-1-phenylethylamine, *rac*-1-phenylethylamine or 2-octylamine), 100 mM NAD^+, TrisHCl buffer pH 8.0 and the respective lysate. For reductive amination enzymatic reactions contained ~30 mM of each substrate (acetophenone or 2-octanone), 100 mM NADH, 0.2 M NH_4Cl, potassium phosphate buffer pH 7.0 and the respective lysate. In general, 50 µL of *S. virginiae* lysate or 50 µL of *E. coli* BL21-AI lysate were used for the reactions in a total volume of 1 mL. The samples were shaken at 1,200 rpm in a 1.5 mL Eppendorf tube at 25°C for 4 h employing an Eppendorf thermomixer. After 4 h of incubation the reactions were quenched by addition of 146 µL of 1M NaOH (to reach pH 9 - 10) and extracted with 2x 250 µL of ethyl acetate. Eventually, the extracts were pooled and further investigated. The mobile phase for TLC consisted of 30% methanol, 70% ethyl acetate and 3 mL of ammonia.

GC method for (*R*)-1-phenylethylamine and acetophenone

Ethyl acetate extracts prepared for TLC were further analyzed by chiral GC. Thereby, about 400 µL of the extracts were mixed with 800 µL ethyl acetate, 20 µL acetic anhydride and 20 µL pyridine. GC conditions (inlets: 250°C, pressure: 0.650 bar, split ratio: 20:1, split flow: 69.3 mL min^{-1}, column: Chirasil DEX-CD 25.0mx320µmx0.25µm nominal, FID detector: flow: H_2: 20 ml min^{-1}, air: 200 ml min^{-1}, make up flow: N_2: 30 ml min^{-1}, temperature: 250°C). Retention times: acetylated (*S*)-1-phenylethylamine: 5.05 min, acetylated (*R*)-1-phenethylamine: 5.49 min; acetylated (*S*)-1-phenylethanol: 1.43 min, acetylated (*R*)-1-phenylethanol: 1.51 min; (*S*)-1-phenylethanol: 1.83 min, (*R*)-1-phenylethanol: 1.77 min; acetophenone: 1.22 min.

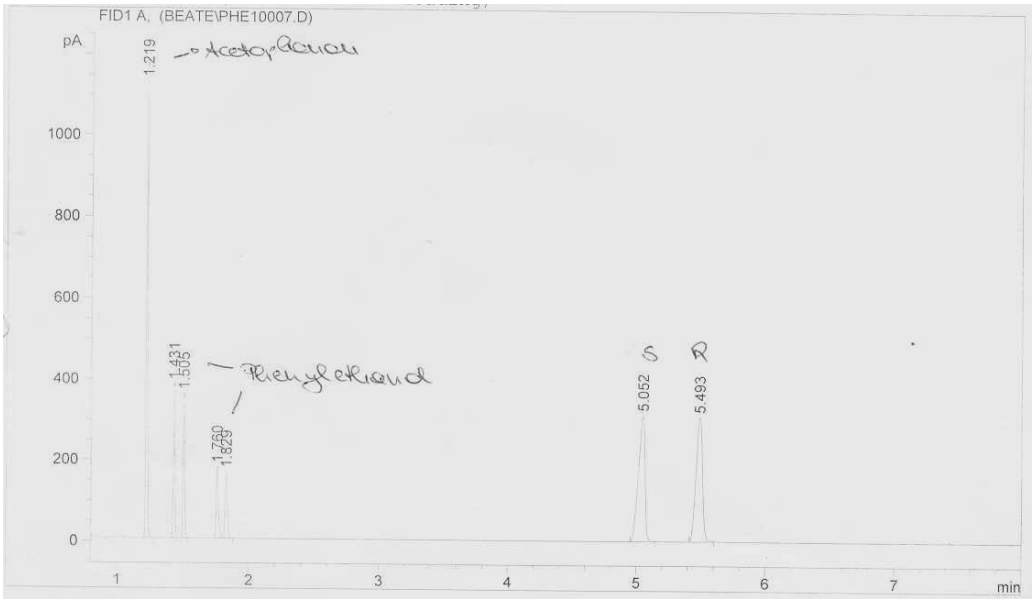
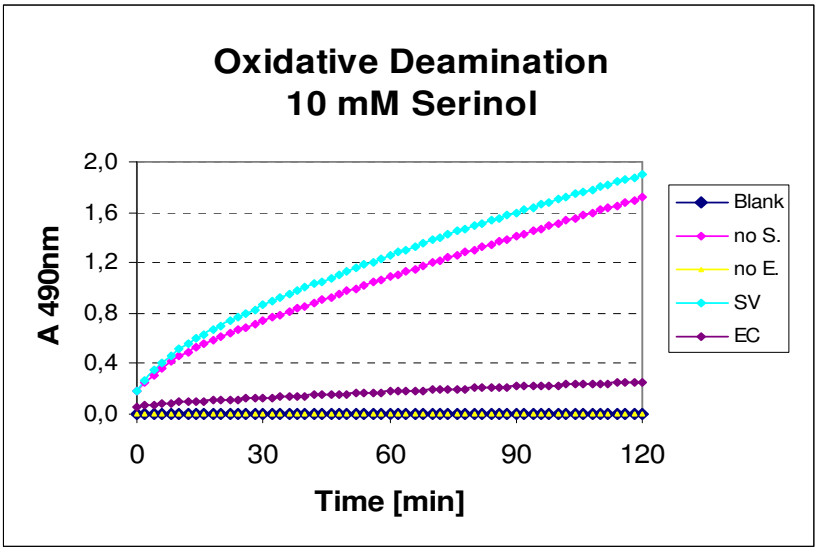


Figure 2.6: GC chromatogram depicting retention times of reference substances.

2.2.4 Experimental results

In literature oxidative deamination activity and reductive amination activity in *S. virginiae* was described. Hence, I tried to reproduce the findings from Itoh and coworkers (Itoh *et al.* 2000) but compared the results with the activity of an *E. coli* lysate.



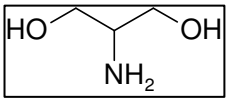


Figure 2.7: Increase in absorption of formazan produced from INT at 490 nm over a 2 h time period. 10 mM serinol was used as substrate. Blank: reaction containing buffer and INT, no S.: reaction containing only the cofactor NAD+ and 50 µL of the *S. virginiae* lysate, no E.: reaction mixture without lysate, SV: reaction mixture containing 50 µL of the *S. virginiae* lysate, EC: reaction mixture containing 50 µL of the *E. coli* BL21-AI lysate as a control.

As expected, in the blank reaction and the reaction mixture without cell free extract no reduction of NAD^+ to NADH occurred. The *S. virginiae* lysate led to a strong increase in formazan due to increased NADH formation compared to *E. coli*. Nevertheless, as depicted in figure 2.7 also the reaction mixture without substrate formed a significant amount of formazan. Thus, increased absorption detected in the reaction containing the cell free extract from *S. virginiae* seemed to be mostly from background reactions. However, there was more formazan dye formed, if substrate was present which could be a result of native cellular alcohol dehydrogenases reacting with serinol.

Subsequently, we investigated the ability of oxidative deamination of (R)-1-phenylethylamine. According to the publication, (R)-1-phenylethylamine revealed a good substrate for the amine dehydrogenase from *S. virginiae*.

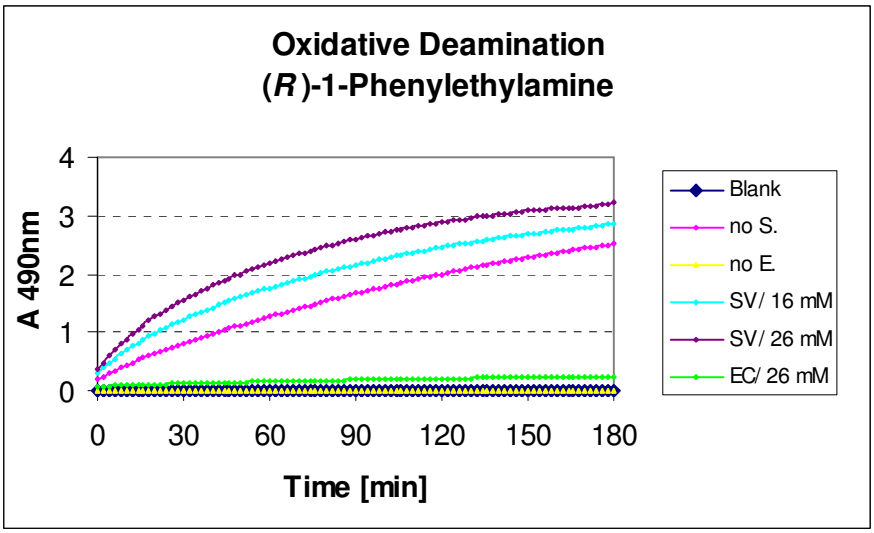


Figure 2.8: Increase in absorption of formazan produced from INT at 490 nm over a 3 h time period. (R)-1-phenylethylamine was used as substrate. Blank: reaction containing buffer and INT, no S.: reaction containing only the cofactor NAD+ and 50 µL of the *S. virginiae* lysate, no E.: reaction mixture without lysate, SV/ 16 mM: reaction mixture containing 50 µL of the *S. virginiae* lysate and 16 mM of (R)-1-phenylethylamine, SV/ 26 mM: reaction mixture containing 50 µL of the *S. virginiae* lysate and 26 mM of (R)-1-phenylethylamine, EC: reaction mixture containing 50 µL of the *E. coli* BL21-AI lysate as a control.

A minimal development of formazan was obtained in the reaction mixture containing *E. coli* BL21-AI. The reaction with *S. virginiae* lysate led to a strong increase in absorption of and higher substrate concentration even increased the activity. Nevertheless, as already observed in the first experiment using serinol, a high background activity was again

measured. As expected, no activity was observed in the blank reaction and the control without lysate, respectively.

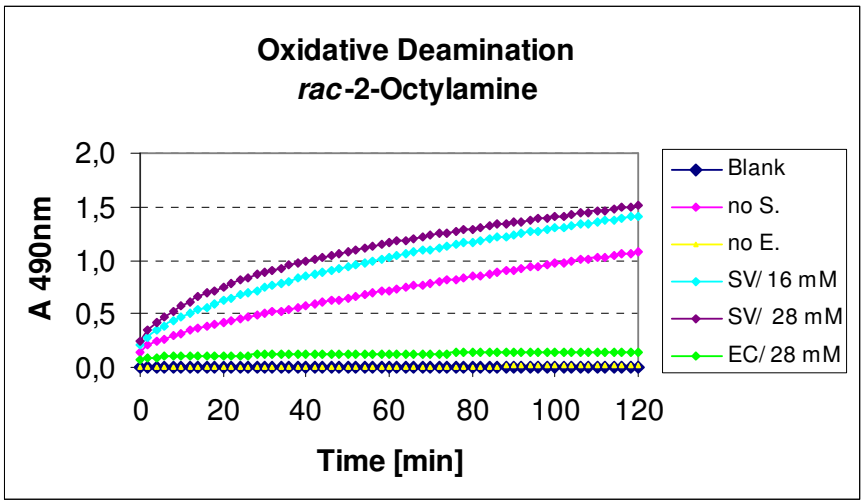
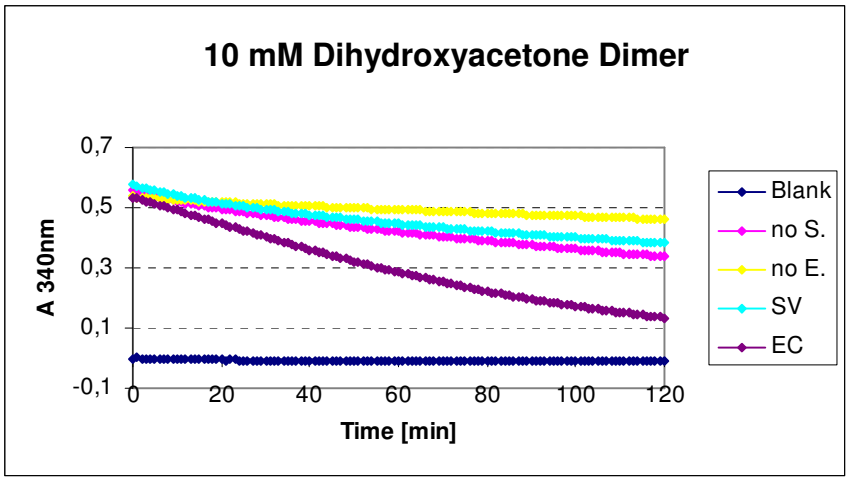


Figure 2.9: Increase in absorption of formazan produced from INT at 490 nm over a 2 h time period. rac-2-octylamine was used as substrate. Blank: reaction containing buffer and INT, no S.: reaction containing only the cofactor NAD+ and 50 µL of the *S. virginiae lysate*, no S.: only cofactor NAD+ was added, no E.: reaction mixture without enzyme solution, SV/ 16 mM: reaction mixture containing 50 µL S. virginiae lysate and 16 mM rac-2-octylamine, SV/ 28 mM: reaction mixture containing 50 µL *S. virginiae* lysate and 28 mM rac-2-octylamine, EC: reaction mixture containing 50 µL *E. coli* BL21-AI lysate and 28 mM rac-2-octylamine as a control.

The next experiment reflected the findings before. No formazan was formed during a time period of 2 h in both control reactions. In the reaction mixture containing *E. coli* BL21-AI lysate hardly any activity was measured. High absorption values were detected again in reactions with *S. virginiae* lysate and the higher the substrate concentration the more formazan was produced. Once again, high background reactions occurred in the *S. virginiae* reaction mixtures without substrate.



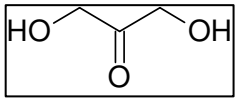


Figure 2.10: NADH depletion over a time period of 2 h using dihydroxyacetone (DHA) as substrate. In this experiment the lysates were 1:5 diluted. Blank: only buffer, no S.: reaction containing the reduced cofactor NADH and 50 µL of the *S. virginiae* lysate, no E.: reaction mixture without lysate, SV: reaction mixture containing 50 µL of the *S. virginiae* lysate, EC: reaction mixture containing 50 µL of the *E. coli* BL21-AI lysate.

For measuring the reductive amination the established NADH depletion assay containing 0.2 M NH_4Cl was used. Surprisingly, this time the highest activity was detected in the reaction mixture containing *E. coli* BL21-AI. Similar absorption values were monitored in the samples containing *S. virginiae* lysate (with and without substrate). Actually, in the reaction mixture without substrate NADH depletion was even slightly higher.

In the next step *S. virginae* lysates were used for further analyses by TLC and GC, respectively, in order to directly detect the desired product.

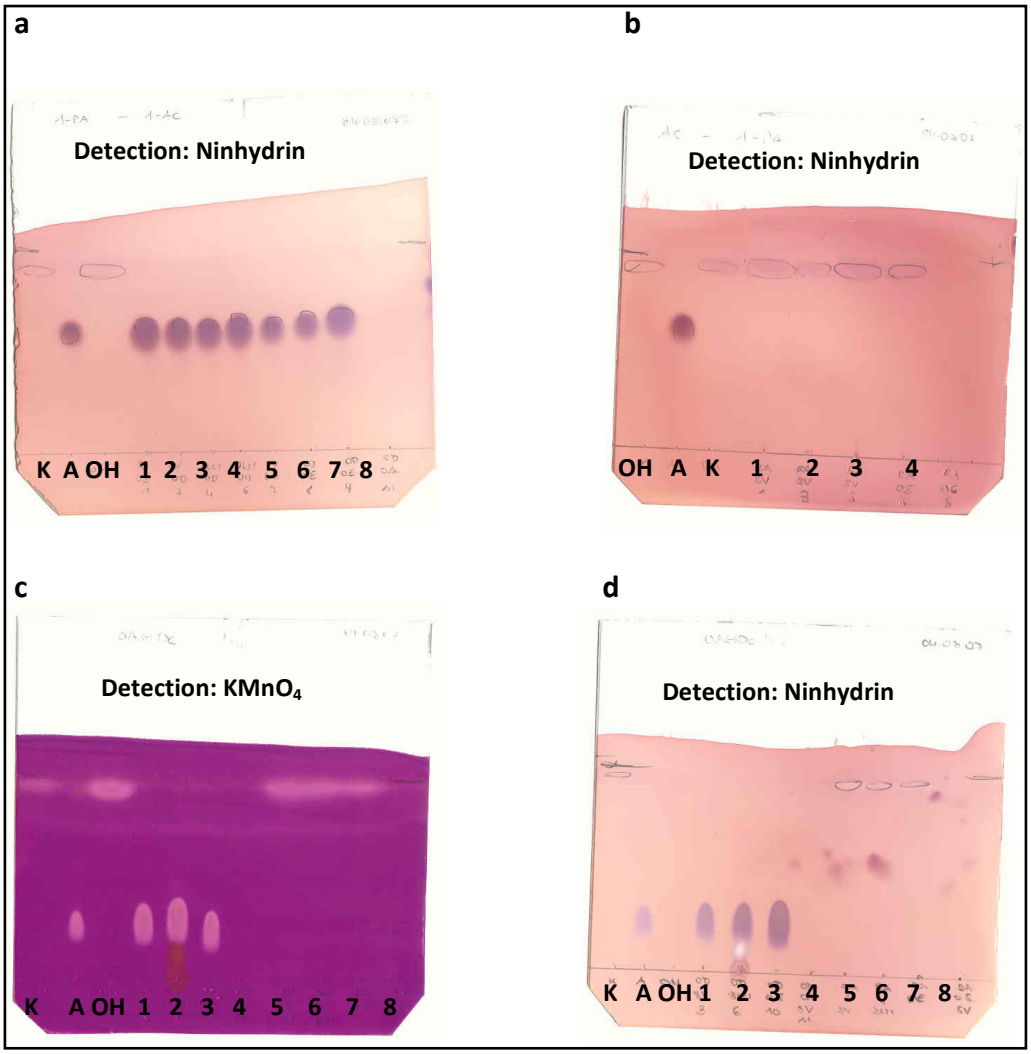


Figure 2.11: TLC Analyses of oxidative deamination and reductive amination. K: ketone, A: amine, OH: alcohol as references. (a) Oxidative deamination reactions using the aromatic model substrate 1-phenylethylamine. References: acetophenone, *rac*-1-phenylethylamin, 1-phenylethanol, 1: *S. virginiae* lysate plus (*R*)-1-phenylethylamine, 2: *S. virginiae* lysate plus *rac*-1-phenylethylamine, 3: *E. coli* BL21-AI lysate plus (*R*)-1-phenylethylamine, 4: *E. coli* BL21-AI lysate plus *rac*-1-phenylethylamine, 5: 200 µL *S. virginiae* lysate plus (*R*)-1-phenylethylamine 6: Reaction control (*R*)-1-phenylethylamine without lysate 7: Reaction control *rac*-1-phenylethylamine without lysate 8: Reaction control without substrate. (b) Reductive amination reactions using the aromatic model substrate acetophenone. References: acetophenone, *rac*-1-phenylethylamin, 1-phenylethanol, 1: *S. virginiae* lysate plus acetophenone, 2: 200 µL *S. virginiae* lysate plus acetophenone, 3: *E. coli* BL21-AI lysate plus acetophenone, 4: Reaction control without lysate 5: Reaction control without substrate. (c) Oxidative deamination and reductive amination reactions using the aliphatic model substrates 2-octylamine and 2-octanone. References: 2-octanone, 2-octanol, *rac*-2-octylamine, 1: *S. virginiae* lysate plus *rac*-2-octylamine, 2: *E. coli* BL21-AI lysate plus *rac*-2-octylamine 3: Reaction control without lysate, 4: Reaction control without substrate, 5: *S. virginiae* lysate plus 2-octanone, 6: *E. coli* BL21-AI lysate plus 2-octanone, 7: Reaction control without lysate, 8: Reaction control without substrate (d) Oxidative deamination and reductive amination reactions using the aliphatic model substrates 2-octylamine and 2-octanone. Controls: 2-octanone, 2-octanol, *rac*-2-octylamine, 1: *S. virginiae* lysate plus *rac*-2-octylamine, 2: *E. coli* BL21-AI lysate plus *rac*-2-octylamine 3: Reaction control without lysate, 4: Reaction control without substrate, 5: *S. virginiae* lysate plus 2-octanone, 6: *E. coli* BL21-AI lysate plus 2-octanone, 7: Reaction control without lysate, 8: Reaction control without substrate.

AMDH activity for oxidative deamination of both model substrates (*R*)-1-phenylethylamine and 2-octylamine, respectively, could not be confirmed using TLC. In addition, the results from TLC also clarified that acetophenone was not converted to 1-phenylethylamine.

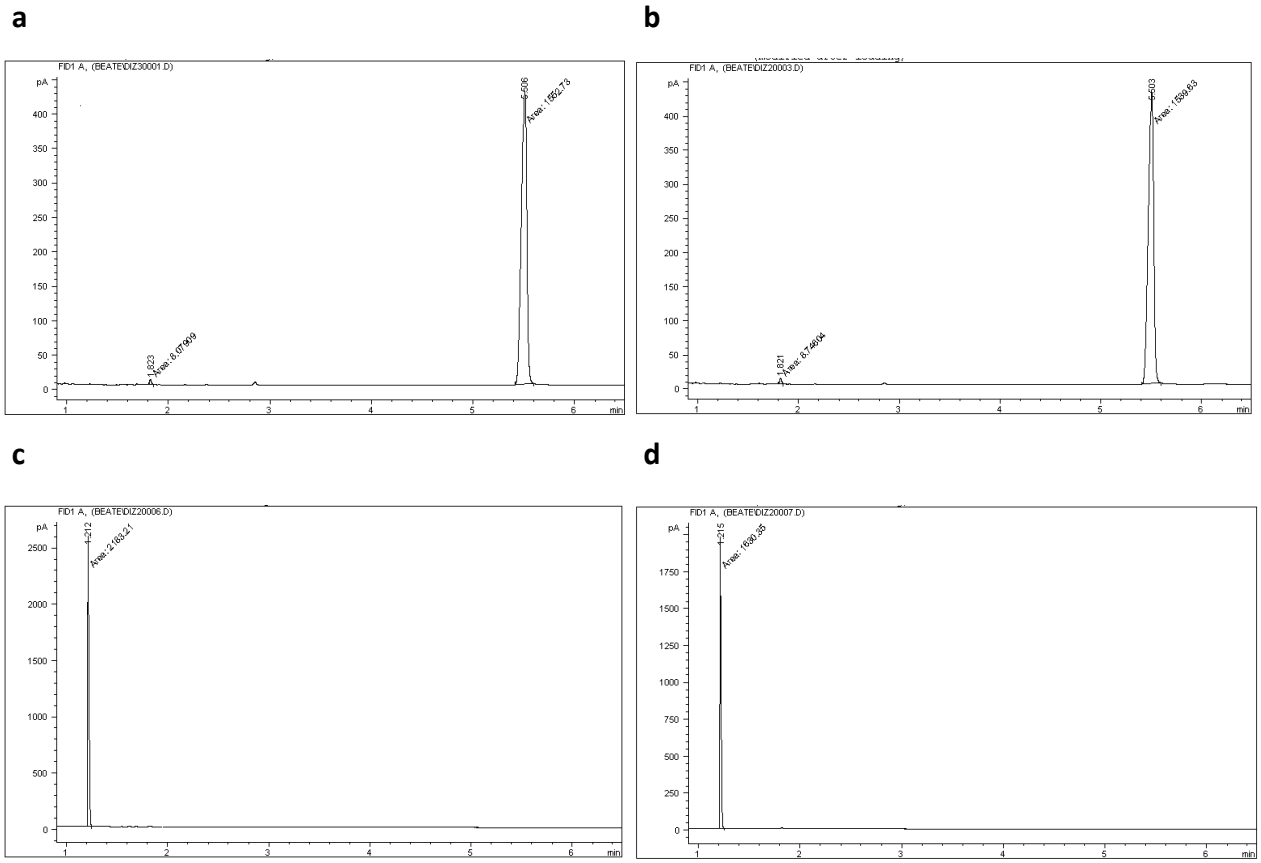


Figure 2.12: Chiral GC analyses – Oxidative deamination (a) Gas chromatogram of an ethyl acetate extract from a reaction mixture containing the *S. virginiae* lysate plus (*R*)-1-phenylethylamine as substrate. (b) Gas chromatogram of an ethyl acetate extract from a reaction mixture containing the *E. coli* BL21-AI lysate plus (*R*)-1-phenylethylamine as substrate. GC Analysis – Reductive amination (c) Gas chromatogram of an ethyl acetate extract from a reaction mixture containing the *S. virginiae* lysate and acetophenone as substrate. (d) Gas chromatogram of an ethyl acetate extract from a reaction mixture containing the *E. coli* BL21-AI lysate and acetophenone as substrate.

Minimal amounts of (*S*)-1-phenylethanol was found in both reaction mixtures for oxidative deamination (Figure 2.12a and Figure 2.12b).

To summarize, GC analysis again confirmed the results obtained from TLC. The model substrates were not aminated or deaminated by *S. virginae* IFO 12827.

2.2.5 Summary and conclusion

Different assays have been established to detect oxidative deamination and reductive amination activity e.g photometric assays (spectrophotometric formazan assay, NADH depletion), TLC and GC, respectively. As described in the above mentioned research article amine dehydrogenase activity in cell free extracts was first measured spectrophotometrically in a DU 800 Spectrophotometer (Beckman Coulter Inc.) using serinol and for the reverse reaction of the amine dehydrogenase the reaction mixture contained dihydroxyacetone. Furthermore, the results of Itoh *et al.* were investigated using the model substrates such as the aromatic substrates *(R)*-1-phenylethylamine, *rac*-1-phenylethylamine and acetophenone or the aliphatic substrates *rac*-2-octylamine and 2-octanone, respectively. The results of the experiments led to the conclusion that controls for the photometric assays were missing in the publication of Itoh *et al.* In this paper there is no experimental evidence that the expected aminated product from dihydroxyacetone was really formed. The strong background due to redox metabolism in cell free extracts was not surprising since dihydroxyacetone generally is a key intermediate of the cellular central metabolism.

Unfortunately, using the established TLC and GC methods, respectively, no evidence of AMDH activity for the reductive amination of the model substrates acetophenone and 2-octanone was adduced under the applied reaction conditions.

2.2.6 References

Baker, P. J., A. P. Turnbull, S. E. Sedelnikova, T. J. Stillman and D. W. Rice (1995) A role for quaternary structure in the substrate specificity of leucine dehydrogenase. *Structure*. **3**, 693-705.

Hummel, W. and M. R. Kula (1989) Dehydrogenase for the synthesis of chiral compounds. *European Journal of Biochemistry*. **184**, 1-13.

Ito, N., A. Matsuyama and Y. Kobayashi (2002-08-13) Amino alcohol dehydrogenase converts keto alcohol to amino alcohol and amino alcohol to keto alcohol. *US6432688 (B1) , DAICEL CHEM (JP)*.

Ito, N., A. Matsuyama and Y. Kobayashi (2003-02-06) Novel amino alcohol dehydrogenase, method for producing said enzyme, and use of said enzyme. *US2003027306 (A1), DAICEL CHEM (US)*.

Itoh, N., C. Yachi and T. Kudome (2000) Determining a novel NAD+-dependent amine dehydrogenase with a broad substrate range from Streptomyces virginiae IFO 12827: purification and characterization. *Journal of Molecular Catalysis B: Enzymatic*. **10**, 281-290.

2.3 Engineering of 2-alkenal reductase from *Arabidopsis thaliana*

2.3.1 Abstract

The 2-alkenal reductase P1 from A*rabidopsis thaliana* shows 'ene' - reductase activity. This enzyme is capable of selectively catalyzing C=C bond reductions of α,β-unsaturated carbonyls. This study describes the engineering of a synthetic variant of P1 from *A. thaliana* (P1syn) and its heterologous expression in *E. coli* TOP10F', followed by enzyme purification and characterization. P1syn was subjected to random and semi-rational mutagenesis strategies to enhance the activity towards the E/Z mixture of the enal 2-[4-methoxy-3-(3-methoxypropoxy) benzylidene]-3-methylbutanal, a new and non-natural substrate. Random mutagenesis by error-prone polymerase chain reaction (epPCR), combination of beneficial mutations, site-saturation mutagenesis and activity screening in the presence of 2% of an organic co-solvent e.g. DMSO, DMF or NMP yielded several P1syn mutants with improved activity. The best variant 2_O7SSMB was constructed from a semi-rational approach combining randomly isolated mutations with a specific amino acid substitution in the active site. Thereby, the stereoselectivity of the enzyme was retained. 2_O7SSMB showed a more than 12-fold increase in activity in the presence of 2% of the organic co-solvents compared to the wild type enzyme. In contrast to P1syn, the bulky substrate 2-[4-methoxy-3-(3-methoxypropoxy) benzylidene]-3-methylbutanal was not converted by the well-known enoate reductase YqjM (Old Yellow Enzyme, OYE) from *Bacillus subtilis* under the applied conditions.

2.3.2 Introduction

An enone reductase (ER) platform with cloned ERs from diverse origins (29 enzymes) were tested in DSM in-house screenings with different substrates and the results were kindly provided by DSM as background information for this study. In the DSM screenings the activities (e.g. NADPH consumption) of the 29 enzymes were compared. Furthermore, the best hits were tested with an additional alcohol dehydrogenase and cofactor recycling, in one pot, under isomerising conditions. The enantiomeric excess (*e.e.*) and the conversion of

2-[4-methoxy-3-(3-methoxypropoxy) benzylidene]-3-methylbutanal (**1a**) to synthon A5 (**1d**) was determined. After the DSM screenings it was concluded that P1 from *A. thaliana* was the best candidate for a further reaction scale up and the most interesting enzyme in the enone reductase platform regarding the bioconversion of 2-[4-methoxy-3-(3-methoxypropoxy) benzylidene]-3-methylbutanal (**1a**). *E. coli* TOP10 containing a synthetic variant of P1 (P1syn) was provided by DSM. P1syn was used as the template for directed evolution.

In literature, our enzyme of interest was initially described as a NADPH:quinone oxidoreductase-like protein in *Arabidopsis* (P1-ζ-crystallin/P1-ZCr) (Babiychuk *et al.* 1995) after chrystallization and preliminary X-ray analysis by Mano and co-workers (Babiychuk *et al.* 1995; Mano *et al.* 2000a; Mano *et al.* 2000b). Based on the activity of the selective reduction of the α,ß-unsaturated bond of 2-alkenals in presence of the aldehyde group, the enzyme was subsequently described as 2-alkenal reductase (also called NADP-dependent oxidoreductase P1, NADPH:2-alkenal α,β-hydrogenase (Mano *et al.* 2002), At-AER (Mano *et al.* 2005) or AtDBR1 (Youn *et al.* 2006). Based on the catalyzed enzymatic reaction, P1 was classified with the EC number 1.3.1.74. The structure of the non-covalent homodimer P1 from *A. thaliana* was already solved and published by Youn and co-workers with a resolution of 2.8 Å (ternary complex I with p-coumaryl aldehyde) in 2006. Furthermore, a multiple sequence alignment with other oxidoreductases and a proposed reaction mechanism of P1 with p-coumaryl aldehyde and 4-hydroxy-(2E)-nonenal (HNE) was included in the publication. The overall fold of P1 belongs to the superfamily of zinc-independent MDRs (medium-chain dehydrogenases/reductases). Each monomer is composed of two domains: a substrate-binding domain (residues 1-137 and 306-345) and a nucleotide binding domain (Rossman fold, residues 138-305). Like other typical MDRs P1 shares a glycin-rich motive in the NADP(H) binding site, with one exception AXXGXXG instead of the more common GXXGXXG (Youn *et al.* 2006). The enzyme is 345 AA in length and has a molecular weight of ~38 kDa (Gene locus name: At5g16970; UniProtKB/Swiss-Prot entry: Q39172). P1 has no non-covalently bound FMN and according to literature the enzyme is not capable of catalyzing the reduction of the C=C double bond of enones such as cyclohex-2-en-1-one or 12-oxo-phytodienoic acid (OPDA). Mano and co-workers claimed that P1 from *A. thaliana* is distinct from enone reductases (Mano *et al.* 2002). According to literature, the key function of P1 from *A. thaliana* lies in detoxification of reactive carbonyls derived from lipid

peroxidation (e.g. HNE) (Mano *et al.* 2002; Mano *et al.* 2005) as well as in preventing cell damage from thiol oxidation (Babiychuk *et al.* 1995). Shortly, 2-alkenal reductase activity from P1 has an essential role in antioxidative defense in plants (Figure 2.13) (Mano *et al.* 2002; Youn *et al.* 2006; Papdi *et al.* 2008; Mano *et al.* 2009).

In summary, P1 was shown capable of reducing diamide/quinone linkages (Mano *et al.* 2000a; Youn *et al.* 2006), recognizes 2-alkenals with C_3-C_9 (including 4-hydroxy-2 alkenals such as HNE and 4-hydroxy-(2E)-hexenal) and selective substrates, which have an α,β-unsaturated enone structure (e.g. 3-buten-2-one, 4-oxo-(2E)-nonenal or 13-oxo-(9E), (11Z)-octadecadienoic acid) (Mano *et al.* 2002; Mano *et al.* 2005). Moreover, Youn et al. reported p-coumaryl aldehyde as the best *in vitro* substrate (Youn *et al.* 2006).

Another MDR with reported alkenal/one reductase activity is the bifunctional leukotriene B4 12-hydroxydehydrogenase/15-oxo-prostaglandin 13-reductase (LTB_4 12-HD/PGR) or also known as NAD(P)H-dependent alkenal/one oxidoreductase (AOR) from rat liver (Dick *et al.* 2001; Dick *et al.* 2004a; Dick and Kensler 2004b). A more recent paper about the LTB4 12-HD/PGR from rat liver was published from Itoh and co-workers (Itoh *et al.* 2008). Very recently, a 2-alkenal reductase from *Artemisia annua* L. was cloned and characterized from Zhang and co-workers (Zhang *et al.* 2009). The 2-alkenal reductase from *A. annua L.* (ACN65116) shows a 67% amino acid sequence identity and 81% similarity with the 2-alkenal reductase P1 from *A. thaliana* (Z49768) [blastp Multiple Alignment - BLOSUM 62].

In this work a protein blast search (Blastp) for protein structural alignment against the P1 sequence was performed. The results are shown in Figure 2.15 and Table 2.1 as well as Table 2.2.

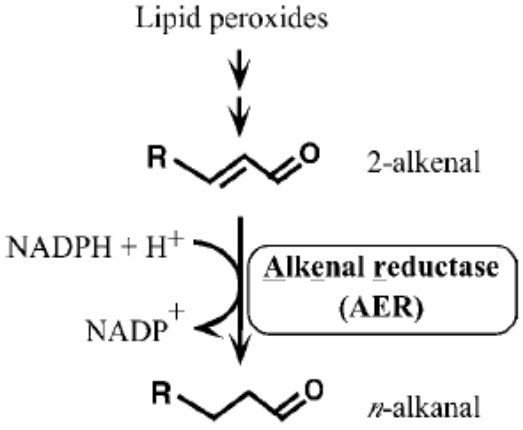


Figure 2.13: Selective Reduction of the α,ß-unsaturated bond of 2-alkenals catalyzed by At-AER (Mano *et al.* 2005)

The well known and highly homologous Old Yellow Enzymes (OYE), a family of flavin containing oxidoreductases, typically catalyze the asymmetric reduction of activated C=C bonds, bearing (at least one) electron-withdrawing substituent(s), such as a carbonyl-or nitro-group (Hall *et al.* 2006). The catalytic mechanism of OYE and its homologs has been studied in great detail and the flavin cofactor plays a key role in the hydride transfer. The catalytic cycle consists of 2 steps, an 'oxidative half reaction' and a 'reductive half reaction'. The electrons are transferred to an electron accepting substrate via FMN (Kohli and Massey 1998; Hall *et al.* 2006; De Wildeman *et al.* 2007; Stuermer *et al.* 2007).

According to literature, the first-reported flavoprotein was initially isolated from brewers' bottom yeast by Warburg and Christian in 1932 and 1933, termed old yellow enzyme, and designated with EC 1.6.99.1 (Karplus *et al.* 1995; Vaz *et al.* 1995; Straßner *et al.* 1999; Kataoka *et al.* 2004). A small group of FMN-dependent oxidoreductases in plants are the 12-oxophytodienate reductases (OPRs), which are homologous to the old yellow enzyme (OYE) from *Saccharomyces cerevisiae* (Breithaupt *et al.* 2009). 12-oxophytodienoate reductase (EC 1.3.1.42) was originally characterized in *Zea mays* L. in 1986 (Vick and Zimmerman 1986) and subsequently 3 isoforms of OPRs have been identified in *A. thaliana* (Williams and Bruce 2002). Molecular cloning and sequence determination of OPR1 from *A. thaliana,* for example, revealed 39% sequence identity to OYE from *S. cerevisiae* (Breithaupt *et al.* 2001). OPR1, OPR2 and OPR3 are the characterized OPRs from *Arabidopsis* and additionally there are 3 as yet uncharacterized putative OPRs, named OPR4, OPR5 and OPR6, with OPR4 and OPR5 being identical (Beynon *et al.* 2009). To summarize, several OPR isoforms have been identified in plants, including 3 isoforms in *Lycopersicon esculentum, 5* in *A. thaliana* and 13 in *Oryza sativa* (Breithaupt *et al.* 2006).

During literature search it became apparent that different EC numbers are associated with the OYE members. To add, enoate reductases are classified as EC 1.3.1.31 (Hall *et al.* 2006; Stuermer *et al.* 2007) and OYE2 from *Saccharomyces cerevisiae* can be found with the EC number 1.5.1.29 in the BRENDA Database.

As already mentioned above, flavin-dependent oxidoreductases typically catalyze the asymmetric reduction of activated C=C bonds. However, FMN is frequently leaking out of the enzyme under different process conditions and therefore an engineered 2-alkenal reductase, which is not depending on FMN poses a worthwhile alternative.

2.3.3 Objectives

In this study, the company DSM was interested in the stereoselective enzymatic reduction of the C=C double of the industrially relevant α,β-unsaturated carbonyl compound 2-[4-methoxy-3-(3-methoxypropoxy) benzylidene]-3-methylbutanal (**1a**). The reaction scheme is depicted in Figure 2.14.

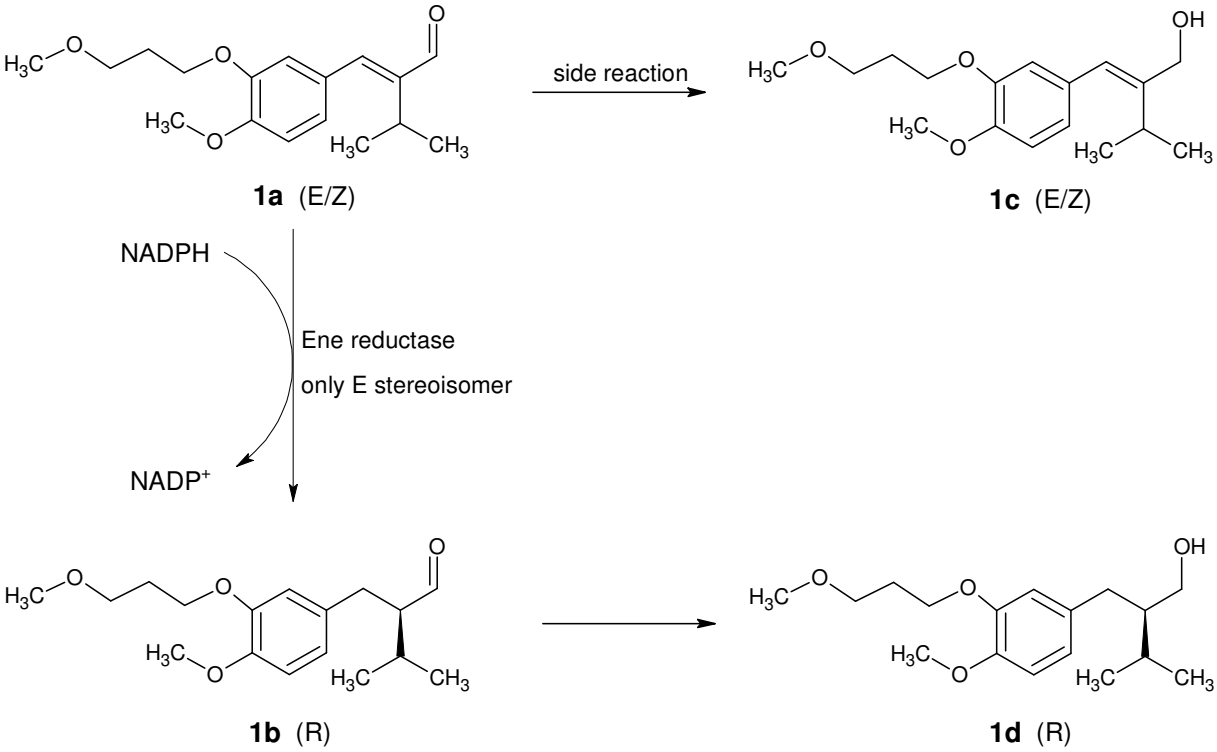


Figure 2.14: Reaction scheme of P1 using cell free extracts. The undesired carbonyl reduction was attributed to *E. coli* reductases present in the cell free extracts. The scheme was made with ACD/ChemSketch 2.0.

The aim of this study described herein was focused on the activity increase towards the non-natural substrate **1a** and improvement of solvent tolerance and solvent stability of P1syn. Directed evolution applying different mutagenesis strategies for engineering of the 2-alkenal reductase P1syn was pursued in this work.

The establishment of a screening system for solvent stability measurements and screening for more stable mutants was part of the diploma thesis “Enhancement of organic solvent stability of an enone reductase” from Thomas Bergner.

2.3.4 Materials and experimental procedures

2.3.4.1 *Materials*

Escherichia coli K12 TOP 10F´ {*lac*I^q, Tn*10*(Tet^R)} *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) ϕ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G (Invitrogen, Carlsbad, CA, USA) was used as a host for basic genetic work and library construction (TOP10F´ Electrocomp™ Kits from Invitrogen). The E/Z enal 2-[4-methoxy-3-(3-methoxypropoxy)benzylidene]-3-methylbutanal (**1a**) as well as the other reference materials for HPLC (saturated aldehyde product (**1b**), E/Z allyl alcohol (**1c**) and synthon A5 (**1d**) were kindly provided by DSM (Geleen, NL). Gene Morph® II Random Mutagenesis Kit and *PfuUltra™* High-Fidelity DNA Polymerase were from Stratagene (La Jolla, CA, USA). All other DNA-modifying enzymes (restriction enzymes, Taq polymerase and T4 DNA Ligase) as well as GeneJET™ Plasmid Miniprep Kits were purchased from *Fermentas* Life Sciences (Burlington, ON, Canada). For gel purification of DNA the QIAquick® gel extraction kit from QIAGEN (Chatsworth, CA, USA) was used. Tryptone as well as yeast extract was purchased from Oxoid Ltd. (Basingstoke, Hampshire, UK) and the antibiotic ampicillin sodium salt was from Sigma Aldrich (St. Louis, MO, USA). Ampicillin was used in a final concentration of 100 µg mL^{-1} throughout this work. Chemicals were ordered from Carl Roth (Karlsruhe, Germany) or Sigma Aldrich (St. Louis, MO, USA). Primers were purchased from IDT (Coralville, IA, USA) or Invitrogen (Carlsbad, CA, USA) and sterile distilled water was from Fresenius Kabi Austria (Graz, Austria). The synthetic double stranded DNA fragment with two AA positions saturated was designed by Sloning (Puchheim, Germany). All the material for SDS-PAGE was from Invitrogen and chemicals used for native PAGE were purchased from Serva (Amstetten, Austria).

2.3.4.2 *Mutant library creation and sequencing*

A synthetic gene coding for P1 (Accession number: Z49768) was kindly provided by DSM. The gene was cloned into the high copy number plasmid pMS470Δ8 (Reiter *et al.* 2000) via *Nde*I and *Hin*dIII restriction sites and transformed into *E. coli* TOP 10F'. The synthetic gene was initially modified by introduction of several mutations by error-prone PCR. 2 rounds of epPCR were performed. The first round of epPCR was performed with the Gene Morph® II Random Mutagenesis Kit. The target DNA was amplified with Mutazyme II DNA Polymerase

in a reaction mixture containing 500 ng of target DNA, 1x Mutazyme II reaction buffer, 0.2 mM of each dNTP, 0.4 µM of forward and reverse primer (Table 2.3), 1 µL of Mutazyme II DNA Polymerase (2.5 U μL^{-1}) and ultra pure water in a final volume of 50 µL. The PCR reaction was conducted with 95°C for 2 min, 30 cycles of 95°C for 30 sec, 60°C for 30 sec, 72°C for 1 min 20 sec and 72°C for 10 min for final elongation. After gel purification, the mutated PCR product was used as a megaprimer to generate the complete plasmid. Conditions of Megaprimer PCR were 50 ng of plasmid DNA, 500 ng of purified mutated PCR products, 1x *PfuUltra™* HF reaction buffer, 0.2 mM of each dNTP, 1.6 µL DMSO, 1 µL of *PfuUltra™* High-Fidelity DNA Polymerase and ultra pure water in a final volume of 50 µL. The PCR program was 95°C for 2 min, 25 cycles of 95°C for 60 sec, 55°C for 1 min, 68°C for 12 min and 72°C for 15 min. The next steps were *Dpn*I digestion and generation of the mutant libraries in *E. coli* TOP 10F‘ by electroporation of the electrocompetent cells with a Gene Pulser II from Bio-Rad (Hercules, CA, USA). After 1 h of incubation at 37°C and 800 rpm, the library was plated onto bio-assay dishes (245x245x25 mm) containing 250 mL LB agar (Lennox) supplemented with ampicillin. Wt P1syn served as positive control, and *E. coli* TOP10F’ harboring the vector control pMS470 was used as negative control in every MTP to facilitate direct comparison of the mutants with the wt during screening. Additionally, wells without inoculation were used as sterile controls in every MTP. For the second round of error-prone PCR a conventional Taq Polymerase was employed. A high $MgCl_2$ concentration (7 mM) as well as increased $MgCl_2$ concentration in addition to unbalanced dNTPs (3 mM supplemented with 0.8 mM dCTP/dTTP) was applied. The reaction mixture (50 µl) was composed of 10 ng of template DNA (pMS470 17_P15), 1x Taq buffer, 0.4 µM of each primer (Table 2.3), 0.3 µL Taq DNA Polymerase (5 U μL^{-1}) and ultra pure water. Thermal cycling parameters were 95°C for 2 min, 30 cycles of 94°C for 40 sec, 60°C for 40 sec, 72°C for 1 min and 72°C for 7 min for final elongation. After *Dpn*I digest, the PCR products were purified and digested with *Nde*I and *Hind*III. The digestion products were purified by gel electrophoresis and the mutated fragments were then ligated into the *Nde*I/*Hind*III digested expression vector pMS470 (over night 16°C). Ligation mixtures were again transformed into *E. coli* electrocompetent TOP 10F‘ cells. Mutations of improved variants were determined by sequencing. Sequencing was performed by AGOWA (Berlin, Germany). Interesting mutations were combined (using the Two Stage PCR Protocol for Site-Directed Mutagenesis–Arbeitsanweisung AA.04) or promising positions were saturated (SlonoMax® library). In

addition, we implemented an independent semi-rational approach on the basis of the published structure 2J3J in the PDB (Youn *et al.* 2006) for saturation of Tyr81 and Tyr290. In this case plasmid DNA of enzyme variant 17_P15 obtained from random mutagenesis was used as template.

2.3.4.3 *Colony picking, cultivation and storage of clones*

96 well and 384 well formats, respectively, were filled with µFill™ microplate reagent dispenser from Bio Tek (Bad Friedrichshall, Germany). At all times, 384 well plates and 96 well plates were incubated without agitation in a 37°C room with increased humidity and in addition the plates were put into a covered glass reservoir to reduce evaporation. *E. coli* colonies were picked with the QPixII picking robot from Genetix (New Milton, UK) into 384 well microplates (polystyrene) from Greiner Bio One (Kremsmünster, Austria) containing 50 µL of LB medium (supplemented with ampicillin) per well. After incubation over night, glycerol was added to the bacterial suspensions to a final concentration of 15% (w/v). The plates were stored at -80°C. For screening, the clones were stamped from the 384 well format into 96 well plates (polystyrene, flat bottom) (Greiner Bio One) with a 96 well pin replicator. The cells were grown over night in these plates containing 150 µL LB medium (Lennox) or 2xTY medium (16 g tryptone, 10 yeast extract and 5 g NaCl) per well. After growth for 16-18 h, the cells were replicated (96 well plates, polystyrene, V-bottom) and incubated for another ~20 h at 30°C with 0.5 mM IPTG. The cells were harvested with 3,000 rpm for 20 min at 4°C in an Eppendorf centrifuge 5810R (swing-bucket rotor A-4-62-MTP) and after a washing step, the pellets were frozen and stored at -20°C (at least for 2-3 h).

2.3.4.4 *Enzyme activity assay – Monitoring NADPH depletion*

The variants were tested on improved activity towards substrate **1a**. For the spectrophotometric microplate NADPH depletion assay, the pelleted cells were lysed with 40 µL of lysis buffer (1 mg mL^{-1} lysozyme, 0.1% Tween20 [for the first round of epPCR 1% Tween20 was used], 50 mM PPB pH 7). Quick resuspension at 1,050 rpm (~ 5-10 min on a TITRAMAX 1000) was performed prior to 1 h of incubation at 30°C. If needed, the samples were diluted with the reaction buffer 50 mM MES-NaOH pH 6 (Mano *et al.* 2005) and after a final centrifugation step (3,500 rpm, 15 min, 4°C) the supernatant was further examined.

Initial rate measurements at 340 nm were performed in the plate reader Spectramax Plus 384 (Molecular Devices, München, Germany). The reaction mixtures consisted of 30 µL cell free extracts (1:1 diluted or non diluted), 150 µL reaction solution (0.5 mM **1a**, 2% co-solvent [DMSO, DMF or NMP], 50 mM MES-NaOH buffer pH 6 and 20 µL 2 mM NADPH (dissolved in reaction buffer). The substrate **1a** was always added to the reaction mixtures dissolved in organic co-solvent to overcome its poor solubility in the aqueous system. Fitness profiles using the established NADPH depletion assay were determined to gain information about the quality of the generated libraries. Furthermore, after protein purification, activities of the wt enzymes (P1syn and YqjM) as well as the best variant 2_O7SSMB towards **1a** and α-methyl-*trans*-cinnamaldehyde **(2a)** were monitored with the NADPH depletion assay. In this case the reaction mixtures consisted of 10 µL enzyme (diluted or non diluted), 170 µL reaction solution (0.5 mM of **1a** or 5 mM of **2a**, 2% solvent [DMSO, DMF or NMP], 50 mM TrisHCl buffer pH 7.5 and 20 µL 2 mM NADPH (dissolved in reaction buffer).

2.3.4.5 *Shake flask cultivations*

50 ml of 2xTY medium supplemented with ampicillin were inoculated with the wt strains (P1syn and YqjM) as well as the most interesting variants from the screenings. The shake flasks were incubated over night at 37°C and an agitation of 120 rpm. After 16-18 h, the 1L main culture shake flasks containing 250 ml 2xTY media supplemented with ampicillin were inoculated with bacterial suspensions to reach an OD_{600} of 0.1. The *E. coli* cells were then grown at 37°C, 120 rpm to an OD_{600} of 0.6 - 0.8. The cells were induced with 0.5 mM IPTG and grown for another ~ 20 h at 37°C and 120 rpm. After 20 h of growth the *E. coli* cells expressing the improved enzyme variants were harvested at 5,000 rpm for 30 min at 4°C (Beckman Coulter JA-10 Fixed-Angle Rotor). After one day storage at -20°C, the pellets were resuspended in 15 ml of 50 mM MES-NaOH reaction buffer pH 6. Cell disruption was carried out by ultrasonication with a subsequent centrifugation step at 40,000 rpm for 45 min at 4°C (Beckman Coulter Optima™LE-80K Ultracentrifuge Rotor Ti70). The supernatant was further examined.

2.3.4.6 *Determination of protein concentration*

Total protein concentrations were measured with the BCA™ Protein Assay Kit (Thermo Scientific, Waltham, MA, USA) with BSA as standard.

2.3.4.7 *SDS-PAGE, BN-PAGE and in-gel activity stains*

Prior to SDS-PAGE, crude or purified protein samples were diluted to the desired concentration in deionized water and mixed with NuPAGE® LDS Sample buffer (4x). Cell free extracts and purified enzymes were subjected to SDS-PAGE (NuPAGE® Novex® Bis-Tris Gels). The SeeBlue® Plus2 Pre-Stained Standard was added prior to the sample lanes. The running buffer was 1x MOPS (200 V for 1 h) and the gels were developed with SimplyBlue™ SafeStain (microwave procedure according to the user guide).

BN-PAGE was performed as prescribed by Leber and Bekerle-Bogner. The PROTEAN II xi cell from BioRad was used to cast the gel with 1 mm thickness and 16x20 cm in size with the following recipe. Stacking gel: 1.35 mL acrylamide (30/0.8), 1.67 mL BN gel buffer (6x), 6.98 mL ddH_20, 50 µL APS (10%), and 10 µL TEMED. Running gel (12%): 12 mL acrylamide (30/0.8), 5 mL BN gel buffer (6x), 13 mL ddH_20, 150 µL APS (10%), and 30 µL TEMED. The BN gel buffer (6x) constisted of 3 M ε-aminocaproic acid, and 0.3 M Bis-Tris HCl (pH 7). Crude cell extracts were diluted in 50 mM MES-NaOH pH 6 to final total protein concentrations of 50 – 100 µg and 30 µL of the extracts were loaded onto the native gel together with 6 µL of loading puffer (150 mM ε-aminocaproic acid, 50% (w/v) glycerol and 0.025% (w/v) Coomassie G 250). Electrophoresis was performed at 4°C at constant current (24 mA per gel) with limited voltage at 1,000 V. After approximately 2.5 h of electrophoresis, the gel was incubated in the activity staining solution which comprised 5 mM of **1a**, 2% DMSO, 4 mM NADPH and 50 mM MES-NaOH pH 6 for 1 h 20 min at 37°C as a gel-filter sandwich in a dark moist chamber. For detection of enzyme activity (NADPH depletion), the native gel was placed onto the UV transilluminator, applying the short wave band filter of the Syngene G:Box HR16 (Cambridge, UK) for 80 ms exposure time (Image acquisition software: GeneSnap). Finally, the native gel was stained with Coomassie Brilliant Blue G-250 (75 mL acetic acid, 500 mL ethanol and 2.5 g Coomassie Brilliant Blue G-250).

2.3.4.8 *HPLC analyses*

Analytical scale reactions were performed in a total volume of 500 µl - 1 mL in Eppendorf tubes (30°C, 400 rpm). The reaction mixtures consisted of 2 mM substrate **1a**, 6 mM NADPH and 2% DMF. For comparison of the improved variants and selectivity determination, every sample was divided. The reactions were stopped after 120 min with DCM or ACN. One part

of the samples were extracted 1:1 with DCM and centrifuged at 13,200 rpm for 2 min. The second part, the ACN samples, were directly used for measuring the conversion. The aqueous phase of the DCM samples was removed and the proteins were in the interphase. The organic phase was transferred and subsequently DCM was evaporated (at room temperature for ~ 2 ½ h). HPLC analyses for conversion were performed using a Zorbax 300 SB-C18 (3.0x150 mm, 3.5 µm) from Agilent (Santa Clara, CA, USA). The mobile phase consisted of a gradient of 0.1% formic acid in water and 0.1% formic acid in acetonitrile (ACN was increased from 20% to 90% within 20 min). Flow rate: 0.45 mL min^{-1}, column temperature: 25°C, injection volume: 10 µl, and detection: UV 230 nm. References and samples were diluted with ACN. Retention times: E/Z enal **1a**, double-peak at 12.2 min; **1b**, 11.7 min; **1c**, 9.2 min; **1d**, 9.8 min. For measuring the conversions with purified enzymes, the reactions were stopped after 0.5 h, 1h, 2h, 3h and 4h of incubation. Selectivity determination of the product was performed in Geleen by DSM. For chiral HPLC the Chiralcel OD-H colum from Daicel was used. After rinsing the column with 99.5/0.5 (% v/v) n-heptane/isobutanol plus 0.1% trifluoroacetic acid for ~ 2 h the column was used for peak separation. The mobile phase consisted of 99.5/0.5 (% v/v) n-heptane/isobutanol and the flow rate was 1.6 ml min^{-1}. Furthermore, the column temperature: 45°C, injection volume: 20 µl, and detection: UV 230 nm was used. References were diluted in the mobile phase. For selectivity determination the samples were resolved in 200 µl DCM and subsequently diluted with 800 µl of 99.5/0.5 (% v/v) n-heptane/isobutanol.

2.3.4.9 *Enzyme purification*

For enzyme purification, the strains were cultivated in 1L 2xTY medium. Heterologous expression and cell harvest of *E. coli* TOP10 F' harboring P1syn or the best enzyme variant 2_O7SSMB was performed like stated in 2.3.4.5 *Shake Flask cultivations*. The pellets were stored at -20°C until further use. The cells were resuspended in 20 ml of the binding buffer A (50 mM Tris HCl pH 7.8). The cells were disrupted by ultrasonication and after a subsequent centrifugation step at 40,000 rpm for 45 min at 4°C (Beckman Coulter Optima™LE-80K Ultracentrifuge Rotor Ti70) the supernatant was filtrated through a 0.22 µm filter. Subsequently, the extract was loaded onto a 20 ml QFF anion-exchange column (Q Sepharose Fast Flow) from GE Healthcare Life Science (Wien, Austria). For enzyme elution buffer B (1 M NaCl, 50 mM Tris HCl pH 7.8) was used. The fractions were analyzed by SDS-

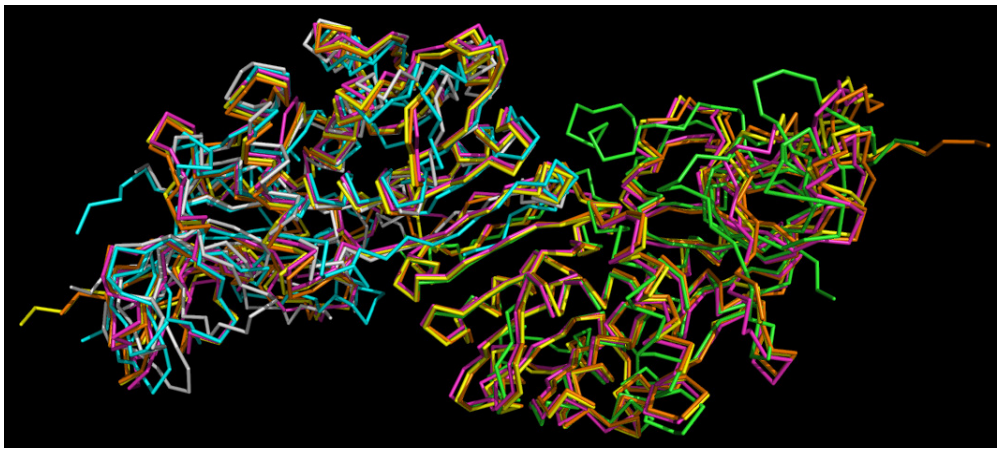
PAGE (Figure 2.33) and the purity was assessed with the Syngene G:Box HR16 using GeneSnap.

2.3.5 Results and discussion

2.3.5.1 *Structural alignment of 2-alkenal reductase P1 with other oxidoreductases*

The results of the alignment of the best hits from a Blastp search against P1 from *A. thaliana* and the results from the 3D-superimposition (PyMol) are shown in Figure 2.15 as well as Table 2.1. and Table 2.2.

a)



b)

Figure 2.15: a) Ribbon representation or b) Cartoon representation of the structural alignment of P1 with oxidoreductases from Blastp search. Chain A of P1 is shown in grey and chain B in green. The colors of the superimposed structures are as follows: 1V3T from *Cavia porcellus* (yellow), 1ZSV from *Homo sapiens* (pink), 1VJ1 from *Mus musculus* (turquoise), 2DM6 *Cavia porcellus* (orange).

Table 2.1: Summary of the results from Blastp and the structural alignment using PyMol

PDB Code	Organism	Enzyme name	Reference	Identity Blastp	Structural alignment RMSD
2J3J	***Arabidopsis thaliana***	**Double Bond Reductase At5g16970**	**(Youn *et al.* 2006)**	**100%**	
1V3T	*Cavia porcellus*	LTB4 12-HD/PGR*	(Hori *et al.* 2004)	42%	2.691
1ZSV	*Homo sapiens*	NADP-dependent leukotriene B_4 12-hydroxydehydrogenase	to be published (Turnbull *et al.* 2005)	38%	2.173
1VJ1	*Mus musculus*	Putative NADPH-dependent oxidoreductase	(Levin *et al.* 2004)	32%	2.219
2DM6	*Cavia porcellus*	LTB_4 12-HD/PGR*	(Hori *et al.* 2004)	42%	2.393

* leukotriene B_4 12-hydroxydehydrogenase/15-oxo-prostaglandin 13-reductase

Table 2.2: (Conserved) active site residues of the superimposed oxidoreductases

PDB Code	(Conserved) active site residues							
2J3J	**Tyr53**	**Tyr81**	**Met138**	**Tyr260**	**Ser287**	**Tyr290**	**B[b]-Ile275**	**B-Tyr276**
1V3T	Tyr49	Gln66	Met124	Tyr245	Tyr273	Gln276	B-Ile261	B-Tyr262
1ZSV	Tyr49	Gln66	Met124	Tyr245	Tyr273	Gln276	B-Ile261	B-Tyr262
1VJ1	Tyr51	Gly77	Met135[a]	structural information missing	Leu290	K293 or F296	[c]	[c]
2DM6	Tyr49	Gln66	Met124	Tyr245	Tyr273	Gln276	B-Ile261	B-Tyr262

[a] in structure with selenium instead of sulfur atom [b] chain B [c] only one chain (1VJ1)

The bifunctional leukotriene B_4 12-hydroxydehydrogenase/15-oxo-prostaglandin 13-reductase represents an $NADP^+$-dependent enzyme responsible for the eicosanoid inactivation (Hori *et al.* 2004). As shown with the results in Table 2.1 with LTB_4 12-HD/PGR the highest identity with P1 from *A. thaliana* was observed. The LTB_4 12-HD/PGR from *C. porcellus* (guinea pig) exhibited 42% identity with the 2-alkenal reductase from *Arabidopsis* (2J3J).

The structural alignment of the ternary complex I of P1 with other oxidoreductases was analyzed to potentially perform structure guided alterations and particularly to define

conserved active site residues. This alignment allowed to decide which residues in the active site or in immediate vicinity to the active site should be targeted for site-saturation mutagenesis to possibly enhance the activity towards the bulky substrate **1a**.

According to literature, Tyr260 plays an important role as a catalytic residue and serves as a general acid/base in the proposed reaction mechanism of P1 (Youn *et al.* 2006). In the course of the structural alignment in this work it was shown that Tyr276 can be aligned with Tyr262 from the other structures. The mentioned residue Tyr262 is conserved among species and was previously described as the catalytic residue for the leukotriene B4 12-hydroxydehydrogenase/15-oxo-prostaglandin 13-reductase (Hori *et al.* 2004). In contrast to Youn *et al*., our analysis indicated that the Tyr276 could maybe account for an alternative acid/base function. The findings in this work seem to be in contradiction to the results from the sequence alignment published by Youn and co-authors (Youn *et al.* 2006).

2.3.5.2 *Heterologous expression of P1syn*

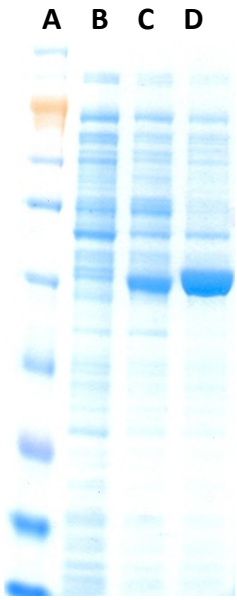


Figure 2.16: Overexpression of P1syn and YqjM was confirmed by SDS-Gel. (A) SeeBlue Prestained Plus2 protein standard (Invitrogen). (B) Cell free extract of *E. coli* TOP 10F' harboring pMS470, (C) Cell free extract of *E. coli* TOP10F' [pMS470 P1syn]. (D) Cell free extract of *E. coli* Top 10F' [pMS470 YqjM]. 10 µg of total protein was loaded.

The wt enzymes P1syn and YqjM were overexpressed in *E. coli* TOP 10F' as soluble proteins (Figure 2.16). The genes encoding P1syn from *A. thaliana* or YqjM from *B. subtilis* were

inserted downstream of the IPTG-inducible *tac* promoter into the expression vector pMS470, resulting in pMS470 P1syn and pMS470 YqjM. SDS-PAGE of the soluble fractions after cultivation and ultrasonication gave a very thick protein band that corresponded to the size of YqjM. Recombinant protein expression of the NADP(H)-dependent oxidoreductase P1syn was also successful. However, overexpression of the plant derived protein was lower under the same cultivation conditions compared to YqjM.

2.3.5.3 *Directed evolution and mutant library screening*

Two different DNA polymerases for error-prone PCR - Mutazyme II and Taq polymerase - were used in this work to create mutant libraries of the NADP(H) dependent 2-alkenal reductase P1syn *in vitro*. Random mutations were introduced into the oxidoreductase gene by PCR under conditions designed to generate an average of 1-3 mutations per gene. Modulation of the mutation rate using Mutazyme II was achieved by varying the amount of target DNA and the number of cycles during PCR. Modulation of the mutation rate using Taq polymerase was accomplished by varying the ratios of nucleotides in the reaction mixture and/or increasing the concentration of $MgCl_2$. Highly electrocompetent *E. coli* TOP 10 F' cells ($1x10^9$) as already mentioned in 2.3.4 *Materials and experimental procedures* were used for transformation of the mutant libraries.

A summary of the sequencing data of all examined variants produced in this study can be found in Table 2.4. Moreover, a complete collection of the relative activities of the improved variants acquired during screening can be extracted from Table 2.5 on page 60 and 61. Primers used throughout this study are listed in Table 2.3.

Mutant libraries were screened for improved activity towards **1a** in the presence of 2% of an organic co-solvent. Co-solvents were used both for better solubility of the highly viscous substrate **1a** and to acquire solvent tolerant improved clones. The solubility of the substrate in buffer was very low in the case of **1a** and therefore a concentration of 0.5 mM was used in the spectrophotometric screening assay. The results of the first round of error-prone PCR were obtained by screening a total of approximately 6800 clones in presence of 2% DMSO. The fitness profile of L2g1 is shown in Figure 2.17. The most improved P1syn variants were confirmed via rescreening and re-rescreening. The relative activities of the 5 best hits from the first round displaying improved activity towards **1a** are depicted in Figure 2.18.

From the first epPCR library we identified two slightly improved mutants. Sequencing analysis confirmed that, 8_M4 had 4 mutations leading to 3 AA substitutions at the positions G80S; D213N; M247L while 24_D18 harbored only a silent mutation at position 15. Investigation of the codon of the silent mutation revealed that TCG is a better codon for serine in *E. coli* compared to TCT in the wt sequence (Kazusa homepage). The three most active mutants from the 1st epPCR round shared an AA substitution at position 57. Comparison of the three following variants 11_D18, 4_E23 and 17_P15 highlighted that the exchange of the arginine to a cysteine at position 57 had the main effect on the enzyme activity. The sequence data is summarized in Table 2.4.

Table 2.3: List of primers used in this study

Primer name	Sequence[a]	Description
P1fw	ATGACTGCGACCAACAAGC	Mutazyme II PCR
P1rv	TTATTCACGCGCGACAACTAC	Mutazyme II PCR
P1*Nde*I	CTAACATATGACTGCGACCAACAAGC	Taq PCR
P1*Hin*dIII	CTAGAAGCTTTTATTCACGCGCGAC	Taq PCR
SSMB_Tyr81_fw	CGATTCAGGGC**NNK**GGTGTATCTCGTATCATC	randomize: Y81
SSMB_Tyr81_rv	GATGATACGAG ATACACC**MNN**GCCCTGAATCG	randomize: Y81
SSMB_Tyr290_fw	CGTAGTCTCCGACTTC**NNK**GATAAATATTCTAA	randomize: Y290
SSMB_Tyr290_rv	CTTAGAATATTTATC**MNN**GAAGTCGGAGACTAC	randomize: Y290
pMSfw	GTGAGCGGATAACAATTTCACACA	insert sequencing
pMSrv	GTTTTATCAGACCGCTTCTGCG	insert sequencing

a) K= T or G; M= A or C, N= unknown (A,T,C or G)

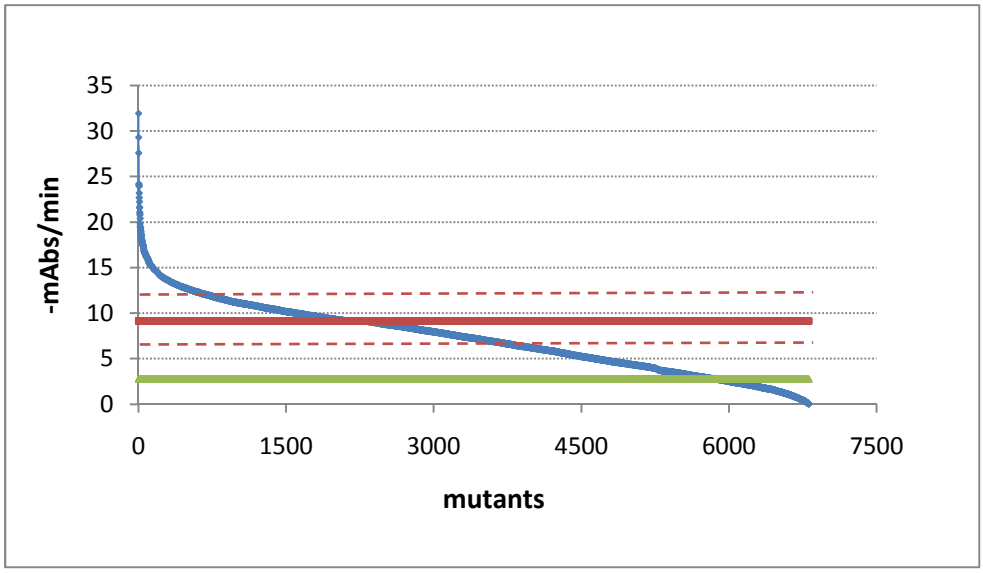


Figure 2.17 Landscape of the first epPCR library (L2g1) created with Mutazyme II. The activity (= decrease in absorption units per min) of the clones is plotted in descent order. The red thick line shows the mean value of the wt and the space between the red dashed lines indicates the standard deviation of the wild type (~19%). The green line depicts the mean value of vector control level (no P1syn). Screening of the library was performed in the presence of 2% DMSO.

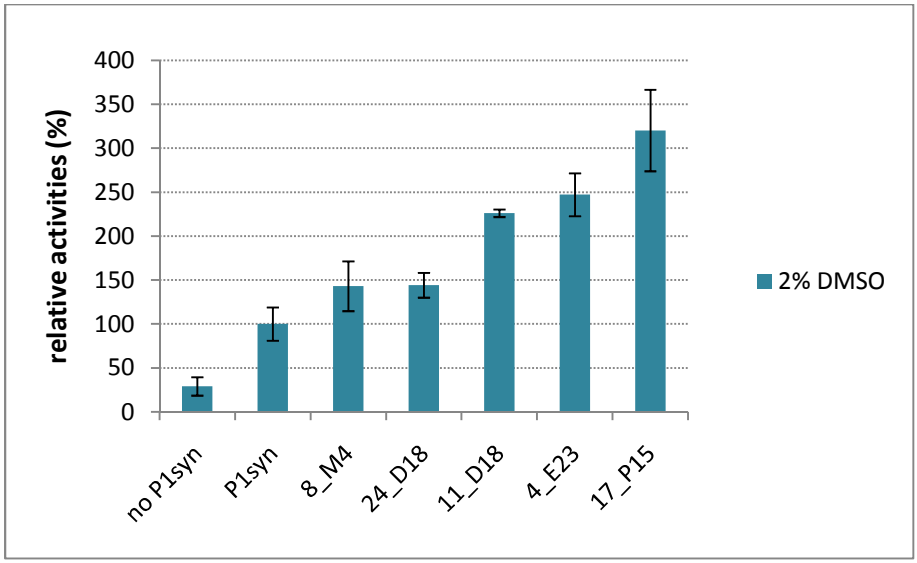


Figure 2.18: Relative activities of the best hits from the first round of epPCR. The activity of the wt (P1syn) was set to 100% and the activities of the improved clones were referred to this value. The absorbance at 340 nm was monitored at 30°C for 4 min with a measurement interval of 12 sec. No P1syn displays the background activity. Error bars represent the mean ± SD (n = 4).

Sequencing of the clone 17_P15 showing the best activity towards **1a** revealed 4 mutations. Three mutations (R57C, F120Y and M115K) led to AA exchanges (Table 2.4). Variant 17_P15 exhibited an about 3-fold improved activity towards the substrate **1a** compared to the wt enzyme.

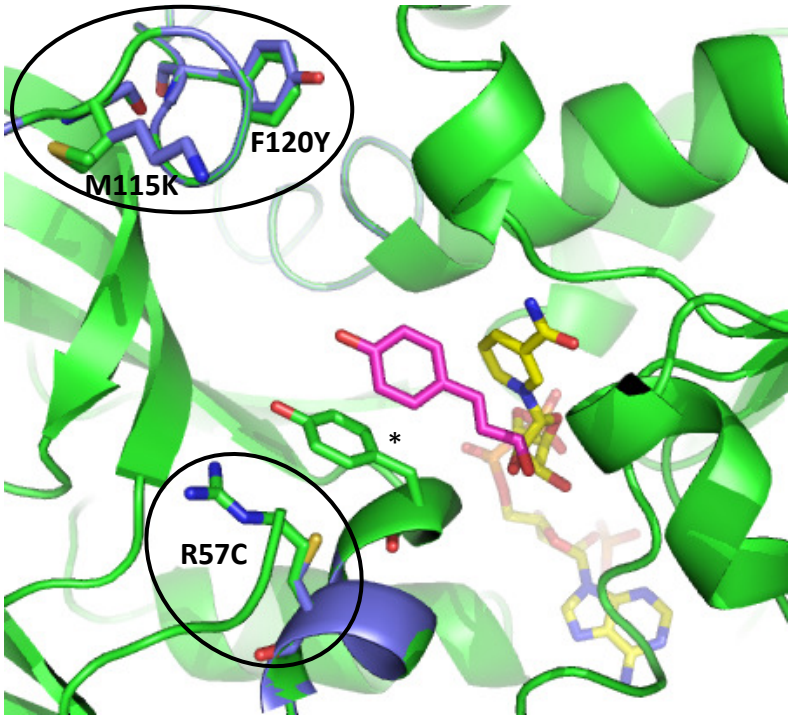


Figure 2.19: Close-up view of the active site of P1 from *A. thaliana* (green) overlaid with 17_P15 (blue). Coordinates were taken from PDB-entry 2J3J. The substrate p-coumaryl aldehyde is shown in pink and NAD^+ is depicted in yellow (Youn *et al.* 2006). The important AA residues of the mutation sites of 17_P15 are circled. R57C had the main effect on the activity. The alignment was generated with PyMol. Tyr53 is marked with an asterisk.

For direct comparison, the structures of the wt and 17_P15 were superimposed using PyMol. All the mutations related to activity enhancement found in the first round of engineering are located in or in immediate vicinity of the active site of the enzyme of interest. In general, the effects of the amino acid exchanges can only be suggested in this work. As aforementioned, R57C had the main effect on the activity (Figure 2.18, Table 2.4). If the substrate **1a**, in particular the phenyl ring of **1a**, is located in a similar way in the active site of 17_P15 as a consequence the cysteine residue might be leading to a higher flexibility of Tyr53 (marked with * in Figure 2.19). Thereby, promoting π-π interactions between the phenolic substrate and the Tyr53 residue (Youn *et al.* 2006).

The AA residues 60-70 are disordered in chain A in the structure of P1 (Youn *et al.* 2006). In the other subunit (chain B) the AA residues of the loop region 60-70 are ordered, but the AA residues have higher B-values. In chain B an interaction between R57 and A69 is shown. Another explanation for the significant activity enhancement caused by R57C might be a conformational change of the loop region 60-70. Thus, a qualitative influence on the activity

of the 2-alkenal reductase variant 17_P15 was observed. M115 and F120, respectively, are at the entrance of the active site and alterations might influence substrate binding. Particularly, with substitution M115K a very hydrophobic AA was exchanged to a charged residue. Thus, this mutation definitely changes the conformation of this area.

To further improve the mutant displaying the highest catalytic activity, a second round of epPCR was applied, using 17_P15 as the starting point. Due to a demand for the industrial application, the solvent tolerance of the enzyme in the presence of different co-solvents was examined. Before starting with the new screening round, the NADPH depletion assay was therefore optimized in the presence of 2% DMF or 2% NMP. In general, by using 2% DMF or 2% NMP, the residual activity of the wild type enzyme compared to the activity achieved in the presence of 2% DMSO was ~32% in the case of 2% DMF and ~24% in the case of 2% NMP, respectively. For a complete comparison, 2% DMSO was used again as a co-solvent in the re-rescreening. In this second epPCR round screening of 3100 colonies (Figure 2.20) yielded 3 prominent variants with increased activity towards **1a** in the presence of 2% of the used co-solvents.

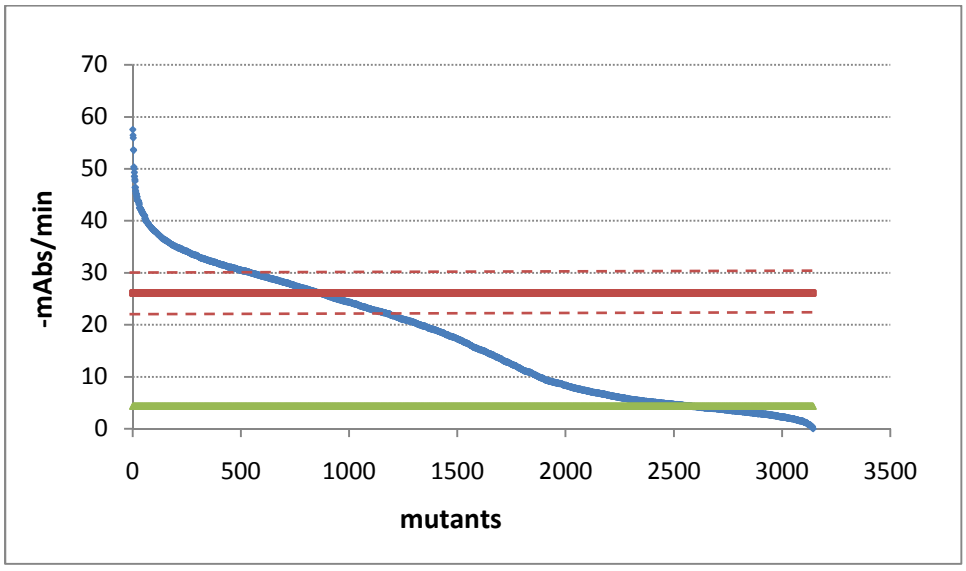


Figure 2.20: Landscape of the second epPCR library (L2g2) created with Taq polymerase. The activity (= decrease in absorption units per min) of the clones is plotted in descent order. The red thick line shows the mean value of the wt and the space between the red colored dashed lines indicates the standard deviation of 17_P15 (~17%). The green line depicts the mean value of vector control level (no P1syn). Screening of the library was performed in the presence of 2% DMF.

Sequencing of the positive clones showed that 3_N12 had one further amino acid substitution M247I as well as 10_K5 and 3_L6 with an exchange M247V and I257V, respectively (Table 2.4). 3_L6 showed a 5-fold enhanced activity in the presence of the co-solvents used compared to the wt (Figure 2.21). Beneficial mutations from the two clones 3_L6 and 10_K5 from the second round were combined using site-directed mutagenesis. This combination led to an ~ 8-fold improvement in activity in the presence of 2% NMP compared to the wt P1syn (Figure 2.21). In contrast, with the combination variant no improvement was observed in the presence of 2% DMSO or 2% DMF compared with 3_L6 or 10_K5, respectively.

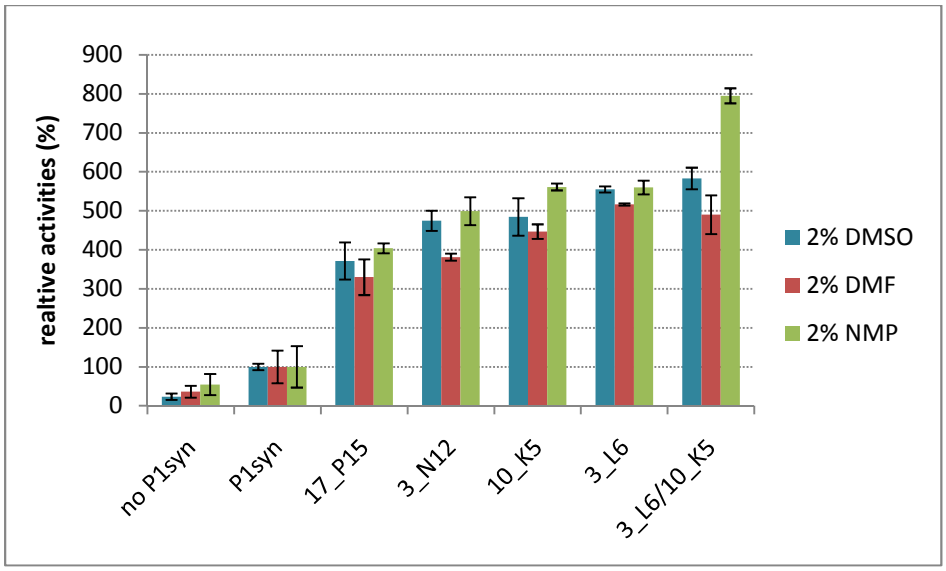


Figure 2.21: Relative activities of the best hits from the 1st and 2nd epPCR and the combination mutant. The activity of the wt (P1syn) was set to 100% and the activities of the improved clones were referred to this value. The absorbance at 340 nm was monitored at 30°C for 4 min with a measurement interval of 12 sec. No P1syn displays the background activity. Error bars represent the mean ± SD (n = 6)

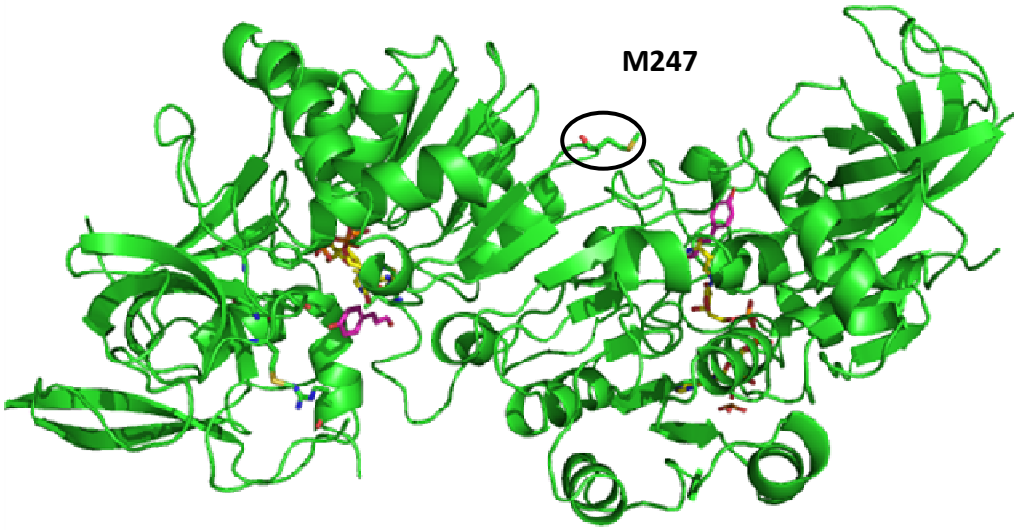


Figure 2.22: Cartoon representation of the overall structure of 2J3J (Youn *et al.* 2006). The position M247 is circled and is located in the dimeric interface in proximity to the substrate entrance channel of chain B. The picture was generated using PyMol.

The methionine at position 247 was found to be substituted 3 times in the course of random mutagenesis (Table 2.4). M247 seems to be a "hot spot" for activity enhancement of P1syn. The structure 2J3J indicates that position 247 is in the dimeric interface on chain A next to the substrate channel of chain B (Figure 2.22). More precisely, the side chain of chain A with M247 interacts with the loop region 60-70 of chain B and this loop may interact with the substrate **1a**. Therefore, one could imagine that an AA exchange at position M247 has an indirect effect on the substrate binding site.

Mutants comprising a mutation at position 247 were subjected to BN-PAGE for investigation if mutations at this respective position might have an influence on the conformational state of the 2-alkenal reductase. The results of this experiment are shown in 2.3.5.3 *Activity staining and BN-PAGE.*

The second interesting position found during random mutagenesis (2nd epPCR) was position 257, which is located in the co-factor binding site (Figure 2.23).

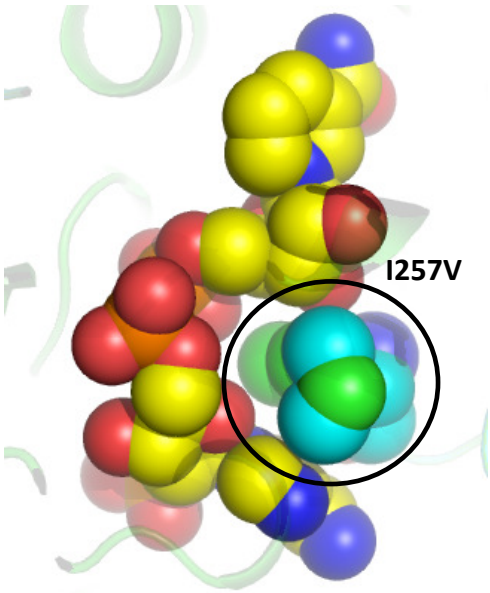


Figure 2.23: Close-up view of NAD^+ and I257V (circled and labeled) which is located in immediate vicinity of the co-factor binding site. Coordinates were taken from PDB-entry 2J3J (Youn *et al.* 2006). The picture was generated with PyMol. Isoleucine is shown with green spheres whereas valine is sown in turquoise.

Figure 2.23 depicts the close proximity of position 257 to the co-factor binding site. The exchange of the isoleucine to a valine in the enzyme variant 3_L6 perhaps leads to an enhanced transfer of the co-factor. Possibly, with the smaller amino acid residue the cofactor's entrance and release is facilitated.

2.3.5.1 *Site-specific randomization*

Only a limited number of amino acid substitutions is accessible by PCR mutagenesis strategies with low error rates (May *et al.* 2000). In an effort to find the best possible amino acid substitution of the specific positions identified by random mutagenesis, we randomized the "wt codons". The starting clone for this approach was the best variant from the first epPCR round, namely 17_P15. A SlonoMax® library was ordered and codons of the selected amino acids M247 and I257, respectively, were substituted with codons for all 20 amino acids, generating a mixture of 20^2 (400) possible 2-alkenal reductase mutants. The "saturated" linear constructs were ligated into the *Nde*I/*Hin*dIII restriction sites of pMS470 and then transformed into *E. coli* TOP10F' via electroporation. Approximately 4000 transformants of the mutein library containing the two simultaneous amino acid substitutions were screened. The landscape of the SlonoMax® library is depicted in Figure 2.24. Sequencing of the plasmid DNA isolated from 24 different colonies revealed that many

of the clones after transformation of the pooled ds DNA fragments possessed deletions or insertions and thereby frame shifts were generated. These deletions or insertions were mainly at the 3' end or 5' end of the gene. In one case a deletion was found in immediate vicinity to the position 257. To summarize, ~ 45% of the sequenced clones were as expected. The landscape achieved from the screening of the SlonoMax® library confirmed both the sequencing results and the expectation to obtain inactive variants by changing residues in the active site.

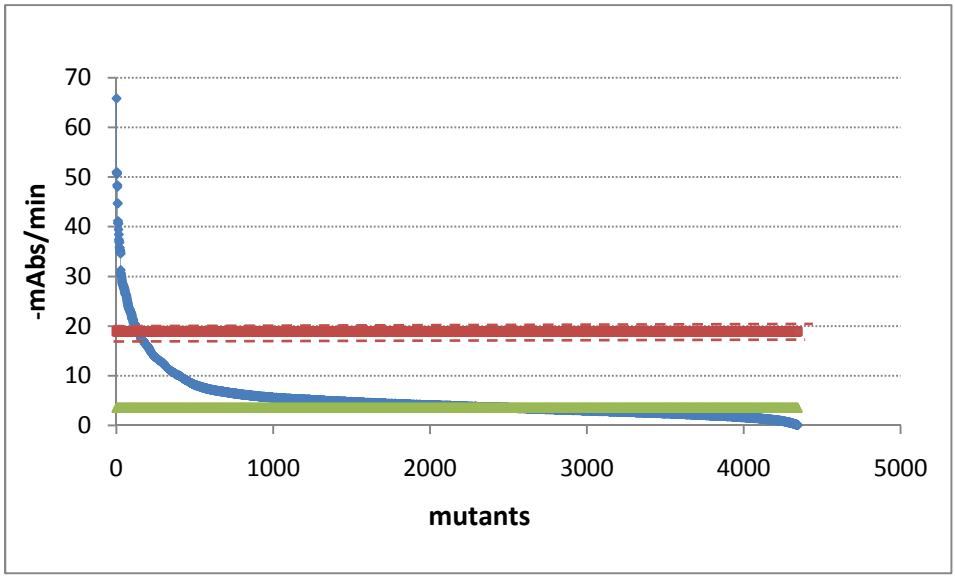


Figure 2.24: Landscape of the SlonoMax® library from Sloning Biotechnology with saturated positions 247 and 257. The activity (= decrease in absorption units per min) of the clones is plotted in descent order. The red thick line shows the mean value of the wt and the space between the red colored dashed lines indicates the standard deviation of 17_P15 (~ 10%). The green line depicts the mean value of the vector control level (no P1syn). Screening of the library was performed in the presence of 2% DMF.

Saturation mutagenesis at the mentioned "hot spots" led to the isolation of two very interesting P1syn variants with increased activity towards the bulky substrate **1a**. The 2 clones 4_C2 and 2_O17 showed a 9 to 10-fold activity enhancement in the case of 2% DMF and a 10-fold improvement in the presence of 2% NMP compared to the wt P1syn (Figure 2.27). The mutants were sequenced and 4_C2 contained the amino acid substitutions M247I and M247T, respectively. Both activity enhanced mutants harbored the substitution I257V. Surprisingly, all sequenced mutants (even slightly improved variants) from the SlonoMax® approach displayed an exchange from isoleucine at position 257 to valine. This exchange was already found in the second round of epPCR. The codon GTC for valine emerged from error-

prone PCR whereas GTT was gained from the SlonoMax® library. This substitution seems to be especially important for activity improvement (Table 2.4. and Table 2.5).

In order to enlarge the space in the active site for the bulky substrate **1a** and to investigate possible synergistic conformational effects by changing two spatially close amino acid positions around the active site (Reetz *et al.* 2005) we followed another semi-rational strategy. In this case plasmid DNA of 17_P15 was again used as template. In the publication of Youn *et al.* an interaction between the phenolic group of Tyr81 and the phenolic group of the substrate p-coumaryl aldehyde is described (Youn *et al.* 2006). Based on the structure of the ternary complex with bound $NADP^+$ and p-coumaryl aldehyde it seemed worthwhile to randomize Tyr81 and Tyr290. If **1a** is located in the active site similarly to p-coumaryl aldehyde, these amino acid exchanges possibly promote enzyme activity by providing more space for the aliphatic side chains of the substrate **1a** (Figure 2.25). The designation of the library was site saturation mutagenesis library B = SSMB. The two codons of interest were replaced by NNK in the primers for the site-saturation mutagenesis (Table 2.3).

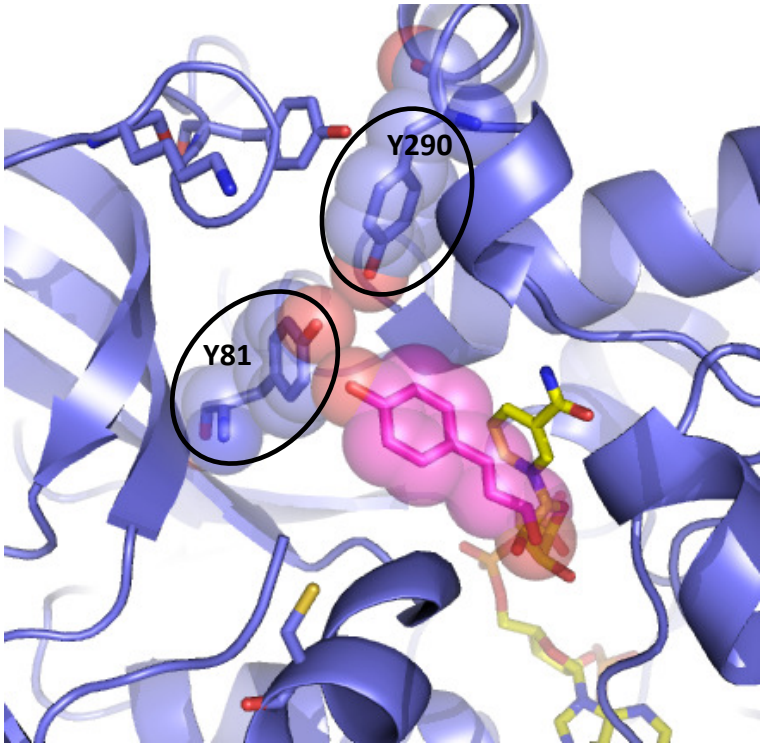


Figure 2.25: Close-up view of the active site of 17_P15 from *A. thaliana.* Coordinates have been taken from PDB-entry 2J3J (Youn *et al.* 2006). The saturation sites of 17_P15 are circled and AA residues are labeled and shown with spheres. The substrate p-coumaryl aldehyde published by Youn *et al.* is shown in pink (spheres) and the NAD^+ is depicted in yellow. The picture was generated with PyMol.

Unfortunately, sequencing of SSMB showed that both positions were not efficiently mutated. More precisely, the AA position 81 was favored over position 290. As a consequence, in this case only 700 colonies were screened for enhanced activity. The landscape of the mutant library SSMB is shown in Figure 2.26.

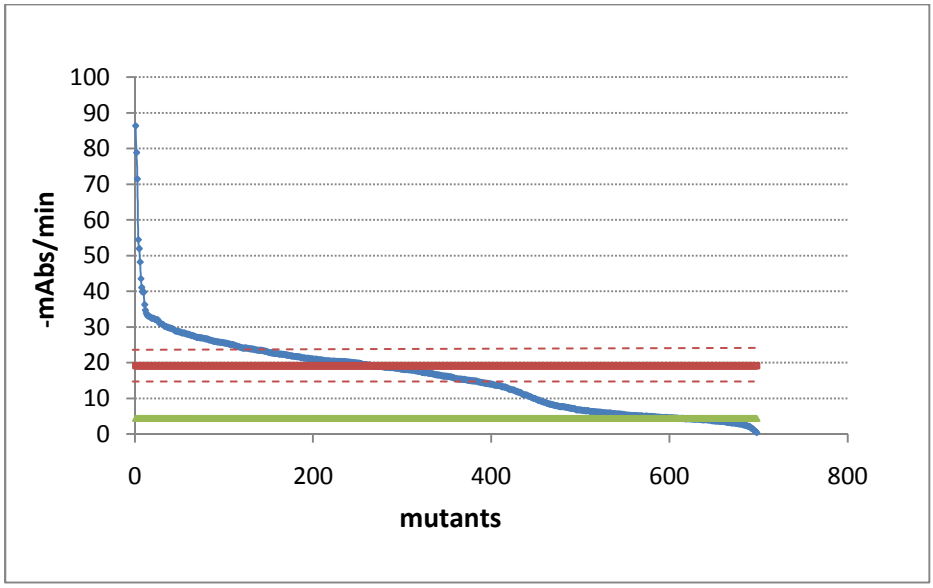


Figure 2.26: Landscape of SSMB. The activity (= decrease in absorption units per min) of the clones is plotted in descent order. The red thick line shows the mean value of the wt and the space between the red colored dashed lines indicates the standard deviation of 17_15 (~ 15%). The green line depicts the mean value of the vector control level (no P1syn). Screening of the library was performed in the presence of 2% DMF.

The most efficient enzyme variant from the semi-rational approaches was 2_O7SSMB. This mutant showed a 12-fold improvement in activity towards **1a** in the presence of 2% DMSO, a 15-fold improvement in activity in the presence of 2% DMF and a 16-fold enhanced activity in the presence of 2% NMP. Surprisingly, this variant included an exchange of Y81 for W. The indole ring possibly leads to a stronger interaction with the substrate **1a**. Only 700 clones were screened in this case and therefore tryptophane is not necessarily the best AA for this particular position. A summary of the relative activities of the best variants from both saturation strategies can be found in the graph of Figure 2.27.

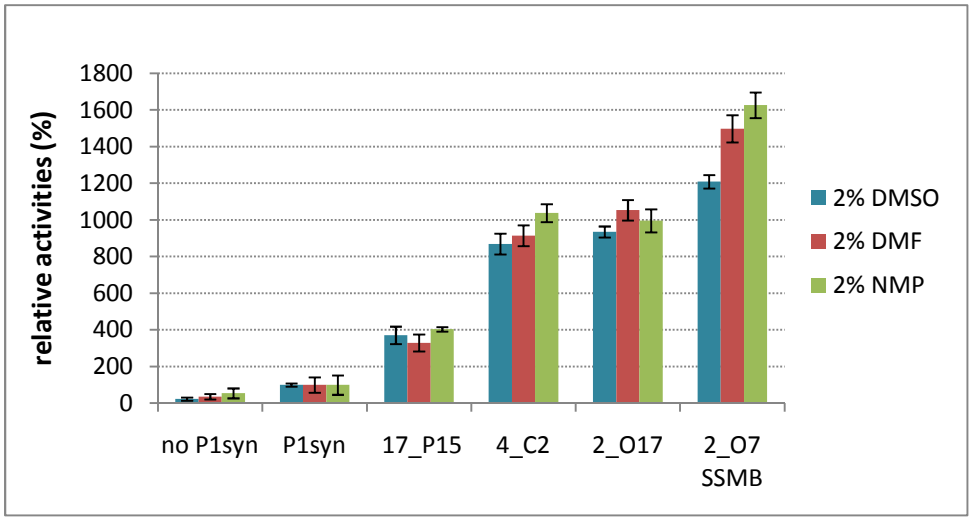


Figure 2.27: Relative activities of the best hits from site-specific randomization. The activity of the wt (P1syn) was set to 100% and the activities of the improved clones were referred to this value. The absorbance at 340 nm was monitored at 30°C for 4 min with a measurement interval of 12 sec. No P1syn displays the background activity. Error bars represent the mean ± SD (n = 6).

Table 2.4: Sequence data of improved variants

Library / generation	mutant	mutation(s)	amino acid exchange	gene sequence
	P1syn	0	0	
L2g1	8_M4	4	G80S; D213N; M247L	GGC → AGC; GAT → AAT; ATG → TTG
L2g1	24_D18	1	0	TCT → TCG
L2g1	11_D18	1	R57C	CGT → TGT
L2g1	4_E23	3	R57C; F120Y	CGT→ TGT; TTC →TAC
L2g1	17_P15	4	R57C; F120Y; M115K	CGT → TGT; TTC →TAC; ATG → AAG
L2g2	3_N12	5	R57C; F120Y; M115K; M247I	CGT → TGT; TTC →TAC; ATG → AAG; ATG→ ATA
L2g2	10_K5	5	R57C; F120Y; M115K; M247V	CGT → TGT; TTC →TAC; ATG → AAG; ATG→ GTG
L2g2	3_L6	5	R57C; F120Y; M115K; I257V	CGT → TGT; TTC →TAC; ATG → AAG; ATC → GTC
Combination	3_L6/10_K5	6	R57C; F120Y; M115K; M247V; I257V	CGT → TGT; TTC →TAC; ATG → AAG; ATG→ GTG; ATC → GTC
SlonoMax®	4_C2	6	R57C; F120Y; M115K; M247I; I257V	CGT → TGT; TTC →TAC; ATG → AAG; ATG→ ATC ; ATC → GTT

SlonoMax®	2_O17	6	R57C; F120Y, M115K; M247T; I257V	CGT → TGT; TTC →TAC; ATG → AAG; ATG→ ACC ; ATC → GTT
SSMB	2_O7SSMB	5	R57C; F120Y, M115K; Y81W	CGT → TGT; TTC →TAC; ATG → AAG; TAC → TGG

Table 2.5: Comparison of P1syn with improved variants using the NADPH microplate assay (1st epPCR round, 2nd epPCR round, SlonoMax® and SSMB)

Library / generation	mutant	MV* relative activity [%]	cosolvent
	P1syn	100 ± 19 [a)]	DMSO
L2g1	8_M4	143 ± 28 [a)]	DMSO
L2g1	24_D18	144 ± 14 [a)]	DMSO
L2g1	11_D18	226 ± 4 [a)]	DMSO
L2g1	4_E23	247 ± 24 [a)]	DMSO
L2g1	17_P15	320 ± 46 [a)]	DMSO
	P1syn	100 ± 8 [b)]	DMSO
L2g1	17_P15	371 ± 48 [b)]	DMSO
L2g2	3_N12	475 ± 26 [b)]	DMSO
L2g2	10_K5	485 ± 48 [b)]	DMSO
L2g2	3_L6	555 ± 8 [b)]	DMSO
Combination	3_L6/10_K5	583 ± 28 [b)]	DMSO
SlonoMax®	4_C2	869 ± 57 [b)]	DMSO
SlonoMax®	2_O17	936 ± 30 [b)]	DMSO
SSMB	2_O7SSMB	1209 ± 37 [b)]	DMSO
	P1syn	100 ± 42 [b)]	DMF
L2g1	17_P15	330 ± 46 [b)]	DMF
L2g2	3_N12	382 ± 9 [b)]	DMF
L2g2	10_K5	447 ± 19 [b)]	DMF
L2g2	3_L6	517 ± 3 [b)]	DMF
Combination	3_L6/10_K5	490 ± 50 [b)]	DMF
SlonoMax®	4_C2	915 ± 57 [b)]	DMF
SlonoMax®	2_O17	1059 ± 56 [b)]	DMF
SSMB	2_O7SSMB	1498 ± 74 [b)]	DMF

	P1syn	100 ± 53 [b)]	NMP
L2g1	17_P15	404 ± 13 [b)]	NMP
L2g2	3_N12	499 ± 36 [b)]	NMP
L2g2	10_K5	561 ± 9 [b)]	NMP
L2g2	3_L6	560 ± 18 [b)]	NMP
Combination	3_L6/10_K5	795 ± 19 [b)]	NMP
SlonoMax®	4_C2	1038 ± 49 [b)]	NMP
SlonoMax®	2_O17	996 ± 63 [b)]	NMP
SSMB	2_O7SSMB	1627 ± 70 [b)]	NMP

***mean value from 4-6 values ± standard deviation, [a)] LB medium/1:1 diluted lysate, [b)] 2xTY medium/non diluted lysate**

2.3.5.2 *Heterologous expression profile of P1syn variants*

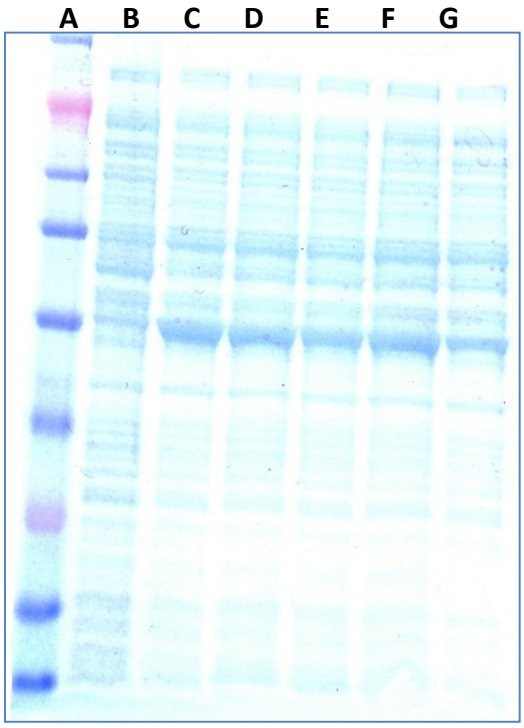


Figure 2.28: Heterologous Expression of P1syn and improved enzyme variants in *E. coli* TOP10F' was confirmed by SDS-Gel. (A) SeeBlue Prestained Plus2 protein standard (Invitrogen) (B) Cell free extract of *E. coli* TOP 10F' harboring pMS470, (C) Cell free extract of TOP10F' [pMS470 P1syn], (D) Cell free extract of *E. coli* Top 10F' [pMS470 17_P15], (E) Cell free extract of *E. coli* Top 10F' [pMS470 4_C2], (F) Cell free extract of *E. coli* Top 10F' [pMS470 2_O17], (G) Cell free extract of *E. coli* Top 10F' [pMS470 2_O7SSMB]. 10 µg of total protein was loaded.

The generated 2-alkenal reductase variants in this work were successfully expressed in *E. coli* TOP 10F'. In our case, activity seemed not to be a consequence of better protein expression, because the expression levels were similar in all clones. However, with the best 2-alkenal reductase mutant 2_O7SSMB the lowest expression was achieved according to the SDS-Gel from Figure 2.28.

2.3.5.3 *Activity staining and BN-PAGE*

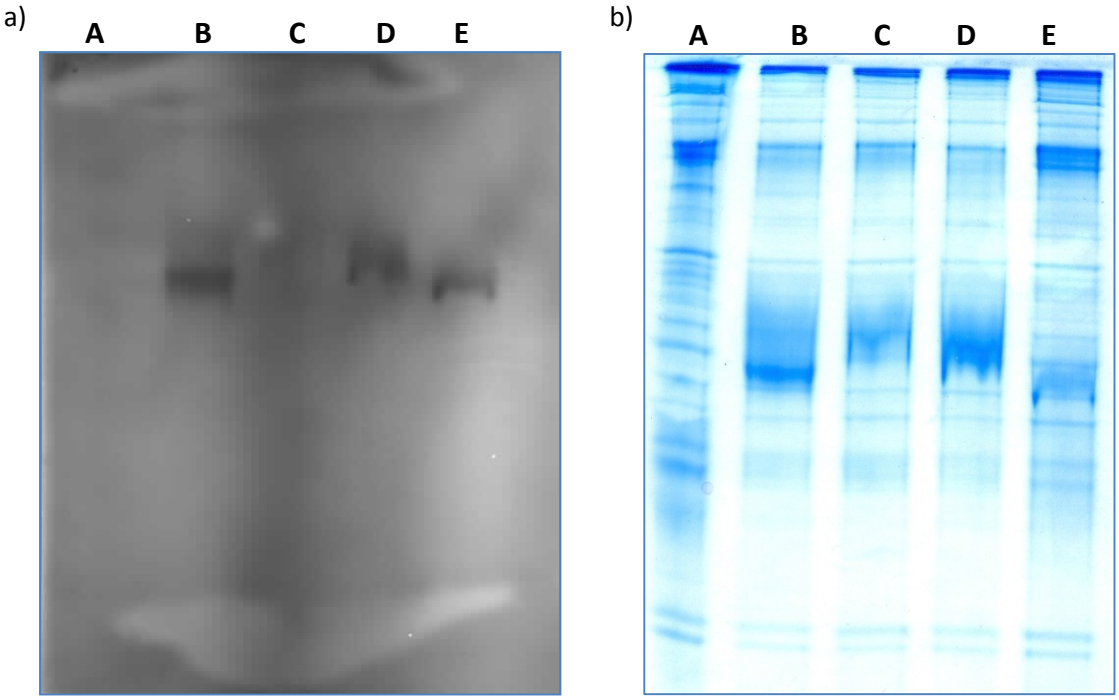


Figure 2.29: a) Activity staining: (A) Cell free extract of *E. coli* TOP 10F' harboring pMS470 (100 µg), (B) Cell free extract of *E. coli* TOP 10F' [pMS470 P1syn] (50 µg), (C) Cell free extract of *E. coli* TOP 10F' [pMS470 4_C2] (50 µg), (D) Cell free extract of *E. coli* TOP 10F' [pMS470 2_O17] (50 µg), (E) Cell free extract of *E. coli* TOP 10 F' [pMS470 10_K5] (100 µg). b) BN-PAGE: (A) Cell free extract of *E. coli* TOP 10F' harboring pMS470 (100 µg), (B) Cell free extract of *E. coli* TOP 10F' [pMS470 P1syn] (50 µg), (C) Cell free extract of *E. coli* TOP 10F' [pMS470 4_C2] (50 µg), (D) Cell free extract of *E. coli* TOP 10F' [pMS470 2_O17] (50 µg), (E) Cell free extract of *E. coli* TOP 10 F' [pMS470 10_K5] (100 µg).

The investigation of the effect of a substitution at position 247 via BN-PAGE and activity staining revealed that this position is not influencing the conformational state of the 2-alkenal reductase. The running performance of the M247 variants and the wt enzyme was very similar as corroborated with Figure 2.29. The protein bands run on nearly the same height which confirms that all the proteins have the same state and no change between monomeric and dimeric forms was observed. In addition according to Youn *et.al.,* the subunit dimerization is achieved mainly through the interactions of the 12 anti-parallel β-sheets across the dimer interface (Youn *et al.* 2006). In-gel activity towards **1a** was detected

in all the bands of the different variants and, as expected, no activity was observed with the vector control.

2.3.5.4 *HPLC analyses – Enzymatic bioreduction of 1a with cell free extracts*

Substrate **1a** was provided by DSM as an E/Z isomeric mixture with an E/Z ratio of 76:24. According to literature (Mano *et al.* 2002; Mano *et al.* 2005) and the information from DSM the E isomer is favored over the Z isomer for double bond reduction by the 2-alkenal reductase from *A. thaliana*. The respective reaction scheme is depicted in Figure 2.14.

Conversions of **1a** using crude cell extracts revealed that the substrate is indeed not completely converted and an amount corresponding to the Z content remains in the reaction mixture. The saturated aldehyde product **1b** is further reduced by *E. coli* background enzymes. The background activity of *E. coli* was tested using cell free extracts of the vector control strain TOP 10F' [pMS470]. As shown in Figure 2.30, no desired product **1b** (retention time: 11.7) was formed with the vector control. Interestingly, the aldehyde group of substrate **1a** is reduced via ADH activity from the *E. coli* background yielding **1c** (retention time: 9.2) (Figure 2.30).

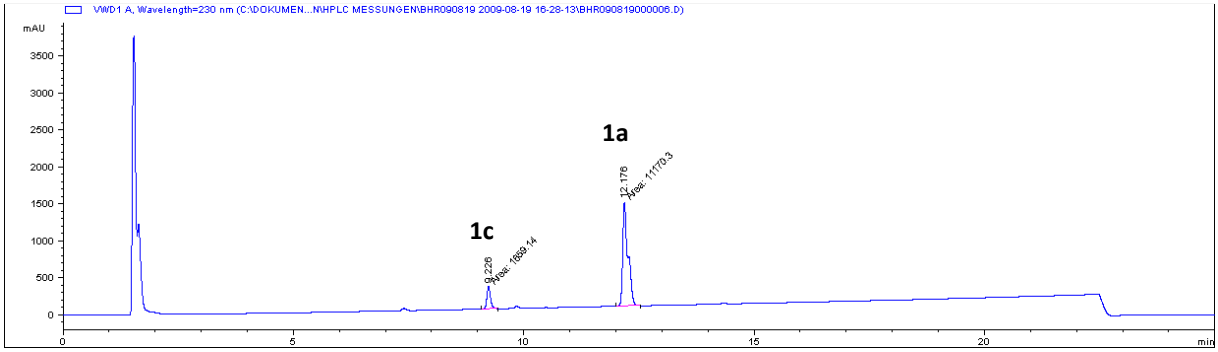


Figure 2.30: Conversion of 1a using a cell free extract of the vector control TOP10F' [pMS470].The incubation parameters were 20 min at 30°C with an agitation of 400 rpm and 2% DMSO as co-solvent.

Subsequently, the desired product (**1b**) was subjected to cell free extracts of the vector control. In this case, the product **1b** was further reduced to **1d** (retention time: 9.8) which is shown in the chromatogram of Figure 2.31.

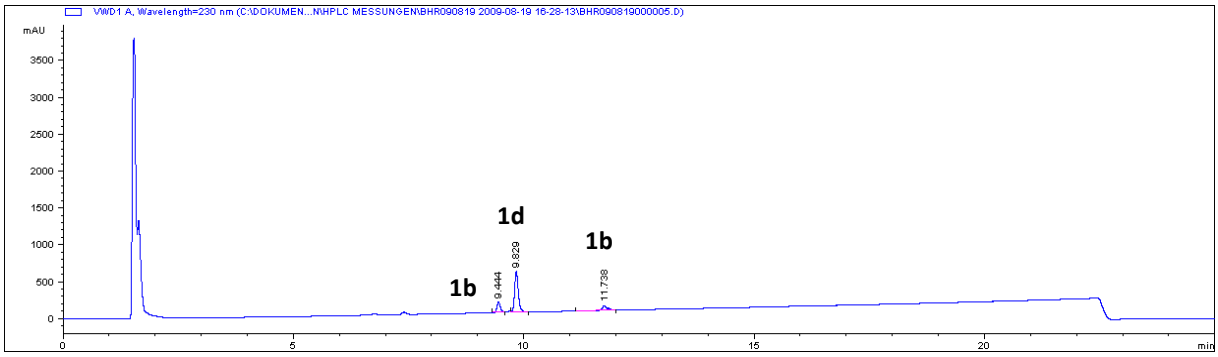


Figure 2.31: Conversion of 1b using a cell free extract of the vector control TOP10F' [pMS470]. The incubation parameters were 20 min at 30°C with an agitation of 400 rpm and 2% DMSO as co-solvent.

For biotransformations with the cell free extracts of the wt enzyme and the variants, the activity of all lysates was adjusted and the reactions were stopped after 2 h of incubation at 30°C and 400 rpm. One part of the samples was stopped with ACN and directly used for measuring the conversion (Figure 2.32). The second part was sent to Geleen for selectivity determination after extraction with DCM. The selectivities are shown in Table 2.6.

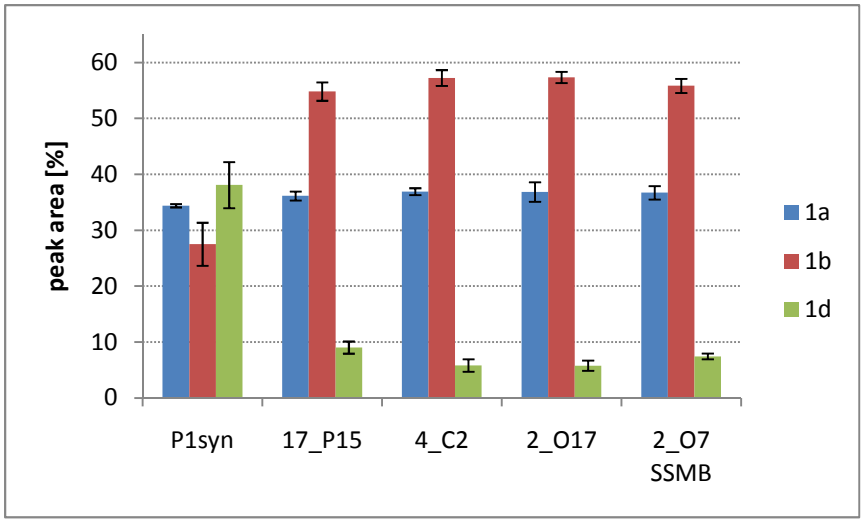


Figure 2.32: Comparison of the conversion of 1a measured with HPLC at 230 nm using cell free extracts of the wt (P1syn) and the activity improved mutants. The values represent the mean ± SD (n=3).

Figure 2.32 shows the product distribution of the biotransformations. The content of **1d** was comparably high in the wt reaction mixture. This can be attributed to the comparably high amount of cell lysate used to obtain the same activity level as for the mutants. Consequently, more background ADH activity was present. On the contrary, the activity improved mutants showed lower amounts of **1d** within the same time. As expected, in the

reaction mixtures of the purified fractions of P1syn and 2_O7SSMB double reduced product **1d** was not detected by HPLC.

Table 2.6: Selectivity measurement and re-calculation of the total *e.e.*

3. Selectivity determination			
variant	***e.e.* 1b (%)**	***e.e.* 1d (%)**	**re-calculated total *e.e.* (%)**
P1syn	66 ± 1	94 ± 1	90
17_P15	86 ± 1	> 95	87
4_C2	90 ± 1	> 95	90
2_O17	90 ± 1	> 95	90
2_O7 SSMB	86 ± 1	> 95	87

The selectivity measurements from DSM revealed that the AA substitutions of the activity improved variants did not negatively influence the stereoselectivity of the enzyme.

2.3.5.5 *Enzyme purification and characterization*

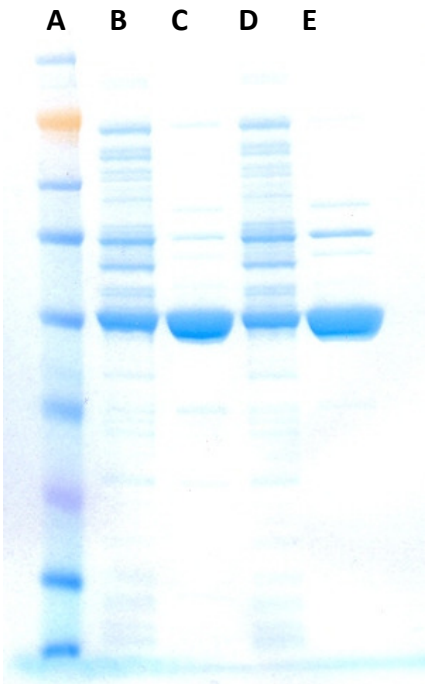
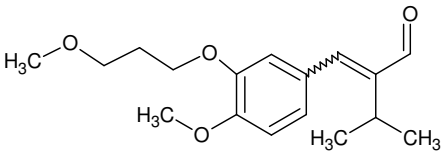


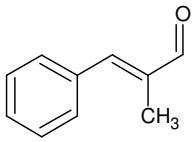
Figure 2.33: SDS-Gel of the wt (P1syn) and 2_O7SSMB after QFF purification. (A) SeeBlue Pre-stained Plus2 protein standard (Invitrogen), (B) Cell free extract of *E. coli* TOP10F' [pMS470 P1syn], (C) Purified fraction of P1syn, (D) Cell free extract of *E. coli* TOP10F' [pMS470 2_O7SSMB], (E) Purified fraction of 2_O7SSMB. The lanes contain 5 µg of protein.

After one step of an anion exchange chromatography, the protein fractions displayed the following purity ~ 81 % for P1syn and ~75 % for 2_O7SSMB, respectively (Figure 2.33). These enzymes were used for characterization in comparison with the OYE member YqjM. The "old yellow enzyme" homologue YqjM from *Bacillus subtilis* was kindly provided by Prof. Peter Macheroux. YqjM from *B. subtilis* was expressed and purified as reported (Kitzing *et al.* 2005).

2.3.5.6 *Specific activities of purified enzymes*



1a (E/Z)



2a (E)

The specific activities of P1syn, 2_O7SSMB and YqjM

Table 2.7: The specific activites of P1syn, 2_O7SSMB and YqjM

	Specific activity [U/mg]					
	Substrate (1a)			**Background activity (no substrate)**		
	2% DMSO	**2% DMF**	**2% NMP**	**2% DMSO**	**2% DMF**	**2% NMP**
P1syn	0.24 ± 0.01	0.05 ± 0	0.04 ± 0	0 ± 0.01	0 ± 0	0 ± 0
2_O7SSMB	3.12 ± 0.02	0.97 ± 0.02	0.57 ± 0.01	0 ± 0.05	0 ± 0	0 ± 0
YqjM	0.18 ± 0.02	0.06 ± 0	0.05 ± 0	0.17 ± 0.05	0.08 ± 0	0.07 ± 0
	Specific activity [U/mg]					
	Substrate (2a)			**Background activity (no substrate)**		
	2% DMSO	**2% DMF**	**2% NMP**	**2% DMSO**	**2% DMF**	**2% NMP**
P1syn	2.20 ± 0.06	0.78 ± 0.10	0.39 ± 0.02	0 ± 0.05	0 ± 0.02	0 ± 0.04
2_O7SSMB	2.18 ± 0.12	0.87 ± 0.11	0.73 ± 0.06	0 ± 0.03	0 ± 0.13	0 ± 0.05
YqjM	2.30 ± 0.02	0.99 ± 0.05	0.75 ± 0.03	0.17 ± 0.03	0.13 ± 0.06	0.10 ± 0.07

Assay conditions: 0.5 mM substrate 1a and 5 mM substrate 2a, 50 mM TrisHCl buffer pH 7.5, 2% DMF, 0.2 mM NADPH, 30°C. Extinction coefficient of NADPH of 6220 M^{-1} cm^{-1} at 340 nm was used in calculation of specific activities. The mean values of 3 measurements are shown in the table.

The activity improved variant 2_O7SSMB exhibited the highest specific activities with **1a** in all three co-solvents used in this study. Notably, with the 2-alkenal reductases, no

background activity was detectable as shown in Table 2.7. In contrast, YqjM showed a high background activity without substrate. This is in accordance to the literature where members of the OYE family are reported to catalyze the oxidation of the nicotinamide cofactor in the absence of substrate (Chaparro-Riggers *et al.* 2007). The reaction mixture containing the OYE YqjM showed a similar activity in presence and in absence of **1a.** The background oxidase activity of YqjM was even slightly higher, especially in the cases of 2% DMF or 2% NMP, respectively. According to the results depicted in Table 2.7 **1a** is not accepted as substrate by YqjM. On the other hand, the activity towards the model substrate **2a** was very similar with P1syn, 2_O7SSMB and YqjM in the case of 2% DMSO. The comparison of the mutant 2_O7SSMB with YqjM in reactions with **2a** displays similar activity towards this substrate in all cases, although the 2_O7SSMB has been evolved for the industrially relevant substrate **1a**. Youn *et al.* claimed that cinnamyl aldehyde is not converted by P1 (Youn *et al.* 2006). Whereby, in this study good conversion was observed with α-methyl-*trans* cinnamaldehyde harboring a methyl group in the alpha position.

To conclude, the specific activity levels demonstrate that the 2-alkenal reductase from A. *thaliana* is clearly the best choice for the bulky substrate **1a.**

2.3.5.7 *HPLC analyses - Enzymatic bioreduction of 1a with purified enzymes*

Biotransformations were performed with purified enzyme fractions in a total volume of 500 µl and the conversion was monitored at different time points by HPLC. No conversion of **1a** was observed with YqjM as demonstrated in Figure 2.34. The results from this experiment confirm the conclusions from the spectrophotometric assay. The best mutant 2_O7SSMB required only 4 h for 62% conversion. In comparison, P1syn reached 23% conversion.

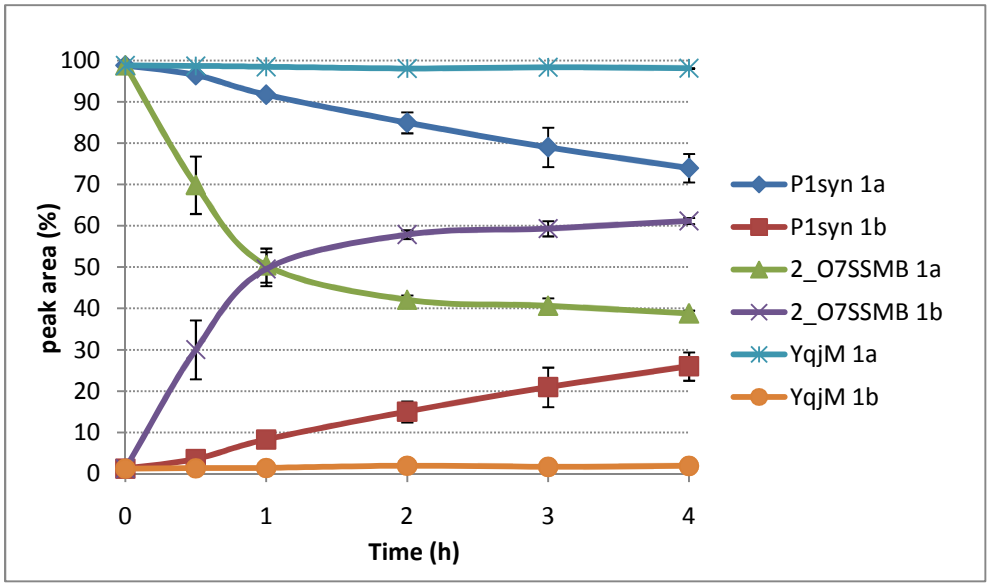


Figure 2.34: Conversion of 1a measured with HPLC at 230nm. The values were determined using area normalization and represent the mean ±SD (n=3).

2.3.6 Conclusions and future prospects

In an attempt to discover the best enzyme with potential for selective reduction of the C=C bond from **1a,** the 2-alkenal reductase P1 from *A. thaliana* was selected. In this work, we combined different methods for directed evolution with further combination of beneficial mutations for successful engineering of P1syn. Initially, 2 rounds of epPCR were employed for activity improvement in the presence of 2% of organic co-solvents (DMSO, DMF and NMP). By combining mutations from random mutagenesis with site-specific randomization methods, a further enhancement of activity could be achieved. The best variant showed a 12-fold improved activity in the presence of 2% DMSO, a 15-fold activity enhancement in the presence of 2% DMF and an even 16-fold improvement in the presence of 2% NMP. Notably, the stereoselectivity was retained.

In this study we confirmed the results from the DSM in-house screenings that P1 displayed a high potential for the reduction of **1a.** Its complementarity to enone/enoate reductases renders the 2-alkenal reductase from *A. thaliana* an interesting target for new industrial applications and further improvement. In fact, the results obtained from this study are very encouraging to screen for activity towards substrates that are different from the usual substrates of enone/enoate reductases.

2.3.7 Acknowledgements

This research was financially supported by DSM, the FFG, the Province of Styria and the SFG. We gratefully thank Stefaan De Wildeman and Natascha Sereinig for their support throughout this study. I wish to thank Beate Pscheidt for support with the structural alignment, Karl Gruber for assistance with the mutant analysis and Margit Winkler for helping with the analytics and for proof reading of this work. I also thank Gerald Mascher and Thomas Bergner for their skillful help in this project.

2.3.8 References

Arnold, F. H. and G. Georgiou (Ed) (2003) *Directed Evolution Library Construction: Methods and Protocols* Humana Press, Totowa, New Yersey

Babiychuk, E., S. Kushnir, E. Belles-Boix, M. Van Montagu and D. Inze (1995) Arabidopsis thaliana NADPH oxidoreductase homologs confer tolerance of yeasts toward the thiol-oxidizing drug diamide. *Journal of Biological Chemistry*. **270**, 26224-26231.

Beynon, E. R., Z. C. Symons, R. G. Jackson, A. Lorenz, E. L. Rylott and N. C. Bruce (2009) The Role of Oxophytodienoate Reductases in the Detoxification of the Explosive 2,4,6-Trinitrotoluene by Arabidopsis. *Plant Physiol.* **151**, 253-261.

Breithaupt, C., R. Kurzbauer, H. Lilie, A. Schaller, J. Strassner, R. Huber, P. Macheroux and T. Clausen (2006) Crystal structure of 12-oxophytodienoate reductase 3 from tomato: Self-inhibition by dimerization. *Proceedings of the National Academy of Sciences of the United States of America*. **103**, 14337-14342.

Breithaupt, C., R. Kurzbauer, F. Schaller, A. Stintzi, A. Schaller, R. Huber, P. Macheroux and T. Clausen (2009) Structural Basis of Substrate Specificity of Plant 12-Oxophytodienoate Reductases. *Journal of Molecular Biology*. **392**, 1266-1277.

Breithaupt, C., J. Strassner, U. Breitinger, R. Huber, P. Macheroux, A. Schaller and T. Clausen (2001) X-Ray Structure of 12-Oxophytodienoate Reductase 1 Provides Structural Insight into Substrate Binding and Specificity within the Family of OYE. *Structure*. **9**, 419-429.

Chaparro-Riggers, J. F., T. A. Rogers, E. Vazquez-Figueroa, K. M. Polizzi and A. S. Bommarius (2007) Comparison of three enoate reductases and their potential use for biotransformations. *Advanced Synthesis and Catalysis*. **349**, 1521-1531.

De Wildeman, S. M. A., T. Sonke, H. E. Schoemaker and O. May (2007) Biocatalytic reductions: From lab curiosity to "first choice". *Accounts of Chemical Research*. **40**, 1260-1266.

Dick, R. A. and T. W. Kensler (2004b) The Catalytic and Kinetic Mechanisms of NADPH-dependent Alkenal/one Oxidoreductase. *Journal of Biological Chemistry*. **279**, 17269-17277.

Dick, R. A., M. K. Kwak, T. R. Sutter and T. W. Kensler (2001) Antioxidative function and substrate specificity of NAD(P)H-dependent alkenal/one oxidoreductase. A new role for leukotriene B4 12-hydroxydehydrogenase/15-oxoprostaglandin 13-reductase. *Journal of Biological Chemistry*. **276**, 40803-40810.

Dick, R. A., X. Yu and T. W. Kensler (2004a) NADPH Alkenal/One Oxidoreductase Activity Determines Sensitivity of Cancer Cells to the Chemotherapeutic Alkylating Agent Irofulven. *Clinical Cancer Research*. **10**, 1492-1499.

Hall, M., B. Hauer, R. Stuermer, W. Kroutil and K. Faber (2006) Asymmetric whole-cell bioreduction of an α,β-unsaturated aldehyde (citral): competing prim-alcohol dehydrogenase and C-C lyase activities. *Tetrahedron Asymmetry*. **17**, 3058-3062.

Hori, T., T. Yokomizo, H. Ago, M. Sugahara, G. Ueno, M. Yamamoto, T. Kumasaka, T. Shimizu and M. Miyano (2004) Structural Basis of Leukotriene B4 12-Hydroxydehydrogenase/15-Oxo-prostaglandin 13-Reductase Catalytic Mechanism and a Possible Src Homology 3 Domain Binding Loop. *Journal of Biological Chemistry*. **279**, 22615-22623.

Itoh, K., K. Yamamoto, M. Adachi, T. Kosaka and Y. Tanaka (2008) Leukotriene B4 12-hydroxydehydrogenase/15-ketoprostaglandin Delta13-reductase (LTB4 12-HD/PGR) responsible for the reduction of a double-bond of the α,ß-unsaturated ketone of an aryl propionic acid non-steroidal anti-inflammatory agent CS-670. *Xenobiotica*. **38**, 249-263.

Karplus, P. A., K. M. Fox and V. Massey (1995) Structure-function relations for old yellow enzyme. *FASEB Journal*. **9**, 1518-1526.

Kataoka, M., A. Kotaka, R. Thiwthong, M. Wada, S. Nakamori and S. Shimizu (2004) Cloning and overexpression of the old yellow enzyme gene of Candida macedoniensis, and its application to the production of a chiral compound. *Journal of Biotechnology*. **114**, 1-9.

Kitzing, K., T. B. Fitzpatrick, C. Wilken, J. Sawa, G. P. Bourenkov, P. Macheroux and T. Clausen (2005) The 1.3 Å crystal structure of the flavoprotein YqjM reveals a novel class of old yellow enzymes. *Journal of Biological Chemistry*. **280**, 27904-27913.

Kohli, R. M. and V. Massey (1998) The Oxidative Half-reaction of Old Yellow Enzyme. *Journal of Biological Chemistry*. **273**, 32763-32770.

Levin, I., R. Schwarzenbacher, D. McMullan, P. Abdubek, E. Ambing, T. Biorac, J. Cambell, J. M. Canaves, H.-J. Chiu, X. Dai, A. M. Deacon, M. DiDonato, M.-A. Elsliger, A. Godzik, C. Grittini, S. K. Grzechnik, E. Hampton, L. Jaroszewski, C. Karlak, H. E. Klock, E. Koesema, A. Kreusch, P. Kuhn, S. A. Lesley, T. M. McPhillips, M. D. Miller, A. Morse, K. Moy, J. Ouyang, R. Page, K. Quijano, R. Reyes, A. Robb, E. Sims, G. Spraggon, R. C. Stevens, H. v. d. Bedem, J. Velasquez, J. Vincent, F. v. Delft, X. Wang, B. West, G. Wolf, Q. Xu, K. O. Hodgson, J. Wooley and I. A. Wilson (2004) Crystal structure of a putative NADPH-dependent oxidoreductase (GI: 18204011) from mouse at 2.10 Å resolution. *Proteins: Structure, Function, and Bioinformatics*. **56**, 629-633.

Mano, J., E. Babiychuk, E. Belles-Boix, J. Hiratake, A. Kimura, D. Inzé, S. Kushnir and K. Asada (2000a) A novel NADPH:diamide oxidoreductase activity in Arabidopsis thaliana P1 ζ-crystallin. *European Journal of Biochemistry*. **267**, 3661-3671.

Mano, J., E. Belles-Boix, E. Babiychuk, D. Inzé , Y. Torii, E. Hiraoka, K. Takimoto, L. Slooten, K. Asada and S. Kushnir (2005) Protection against Photooxidative Injury of Tobacco Leaves by 2-Alkenal Reductase. Detoxication of Lipid Peroxide-Derived Reactive Carbonyls. *Plant Physiology*. **139**, 1773-1783.

Mano, J., F. Miyatake, E. Hiraoka and M. Tamoi (2009) Evaluation of the toxicity of stress-related aldehydes to photosynthesis in chloroplasts. *Planta*. **230**, 639-648.

Mano, J., Y. Torii, S.-i. Hayashi, K. Takimoto, K. Matsui, K. Nakamura, D. Inzé, E. Babiychuk, S. Kushnir and K. Asada (2002) The NADPH:Quinone Oxidoreductase P1- ζ-crystallin in

Arabidopsis Catalyzes the α, ß-Hydrogenation of 2-Alkenals: Detoxication of the Lipid Peroxide-Derived Reactive Aldehydes. *Plant Cell Physiol.* **43**, 1445-1455.

Mano, J., H.-J. Yoon, K. Asada, E. Babiychuk, D. Inzé and B. Mikami (2000b) Crystallization and preliminary X-ray crystallographic analysis of NADPH: azodicarbonyl/quinone oxidoreductase, a plant ζ-crystallin. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology*. **1480**, 374-376.

May, O., P. T. Nguyen and F. H. Arnold (2000) Inverting enantioselectivity by directed evolution of hydantoinase for improved production of L-methionine. *Nature Biotechnology*. **18**, 317-320.

Papdi, C., E. Abraham, M. P. Joseph, C. Popescu, C. Koncz and L. Szabados (2008) Functional Identification of Arabidopsis Stress Regulatory Genes Using the Controlled cDNA Overexpression System. *Plant Physiol.* **147**, 528-542.

Reetz, M. T., M. Bocola, J. D. Carballeira, D. Zha and A. Vogel (2005) Expanding the Range of Substrate Acceptance of Enzymes: Combinatorial Active-Site Saturation Test. *Angewandte Chemie*. **117**, 4264-4268.

Reiter, B., A. Glieder, D. Talker and H. Schwab (2000) Cloning and characterization of EstC from *Burkholderia gladioli*, a novel-type esterase related to plant enzymes. *Appl. Microbiol. Biotechnol.* **54**, 778-785.

Straßner, J., A. Fürholz, P. Macheroux, N. Amrhein and A. Schaller (1999) A homolog of old yellow enzyme in tomato. Spectral properties and substrate specificity of the recombinant protein. *Journal of Biological Chemistry*. **274**, 35067-35073.

Stuermer, R., B. Hauer, M. Hall and K. Faber (2007) Asymmetric bioreduction of activated C=C bonds using enoate reductases from the old yellow enzyme family. *Current Opinion in Chemical Biology*. **11**, 203-213.

Turnbull, A. P., C. Johansson, P. Savitsky, K. Guo, A. Edwards, C. Arrowsmith, M. Sundstrom, F. von Delft and U. Oppermann (2005) Crystal structure of human NADP-dependent leukotriene B4 12-hydroxydehydrogenase. *To be published*.

Vaz, A. D. N., S. Chakraborty and V. Massey (1995) Old yellow enzyme: Aromatization of cyclic enones and the mechanism of a novel dismutation reaction. *Biochemistry*. **34**, 4246-4256.

Vick, B. A. and D. C. Zimmerman (1986) Characterization of 12-Oxo-Phytodienoic Acid Reductase in Corn. *Plant Physiol.* **80**, 202-205.

Williams, R. E. and N. C. Bruce (2002) 'New uses for an old enzyme' - The Old Yellow Enzyme family of flavoenzymes. *Microbiology*. **148**, 1607-1614.

Youn, B., S. J. Kim, S. G. A. Moinuddin, C. Lee, D. L. Bedgar, A. R. Harper, L. B. Davin, N. G. Lewis and C. Kang (2006) Mechanistic and structural studies of apoform, binary, and ternary complexes of the Arabidopsis alkenal double bond reductase At5g16970. *Journal of Biological Chemistry*. **281**, 40076-40088.

Zhang, Y., K. H. Teoh, D. W. Reed and P. S. Covello (2009) Molecular cloning and characterization of Dbr1, a 2-alkenal reductase from *Artemisia annua*. *Botany*. **87**, 643-649.

3 Appendix

3.1 Strains and plasmids

Part 1: Transposon Integration mediated Mutagenesis (TIM)

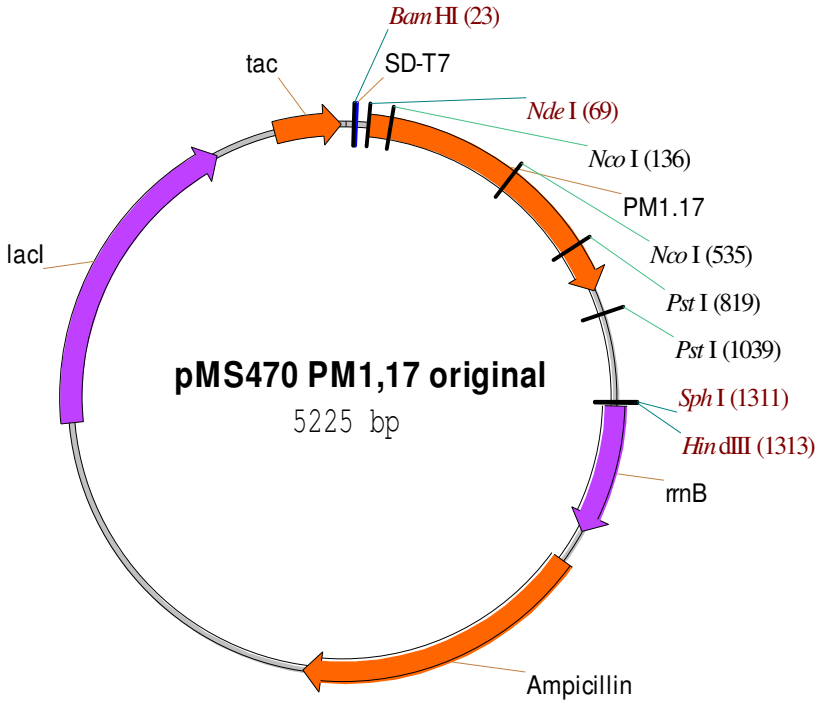
Organism	Plasmid/Clone	VNTI file name	Culture Collection #
Escherichia coli TOP 10F'	[TIM1][a] TIM11H1[b]	TIM1.gb	3044
Escherichia coli TOP 10F'	[TIM2][a] TIM estC+30[b]	TIM2.gb	3045

[a] Designation in the publication [b] Designation in the lab journal

Vector pMS470 PM1,17 original (with EstC from *Burkholderia gladioli*) = pMSPM1,17

Culture collection # 1725

Designation in the publication: pMSEstC



Part 2: Enzymatic reductive amination

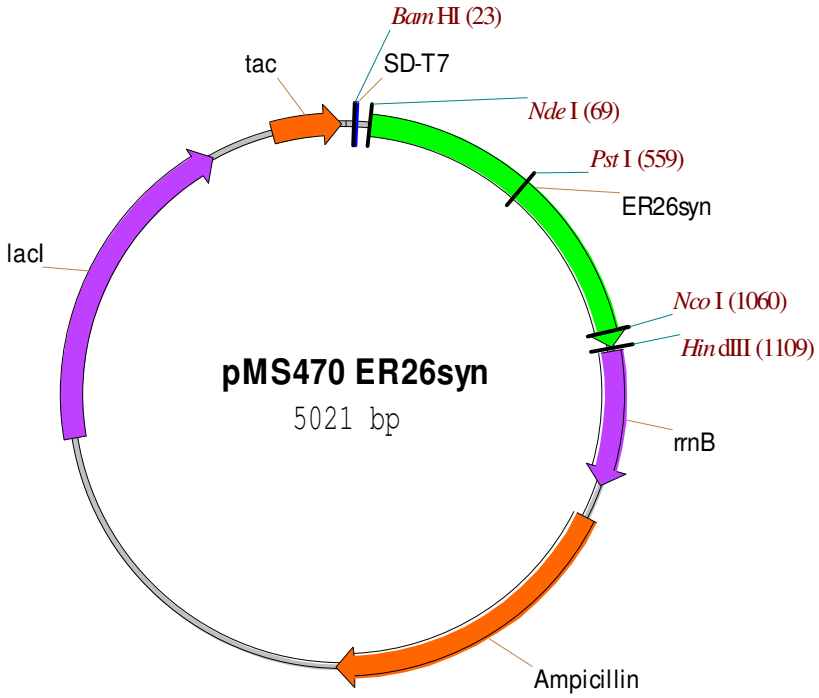
Organism	Culture Collection #
Streptomyces virginiae	5107

Part 3: 2-alkenal reductase

Organism	Plasmid/Clone	VNTI file name	Culture Collection #
Escherichia coli TOP 10F'	[pMS470]	pMS470.gb	3046
Escherichia coli TOP 10F'	[pMS470 P1syn][a] [pMS470 ER26syn][b]	pMS470 ER26syn.gb	5947
Escherichia coli TOP 10F'	[pMS470 YqjM][a] [pMS470 ER24][b]	pMS470 YqjM.gb	3047
Escherichia coli TOP 10F'	[pMS470 17_P15]	pMS470 17_P15.gb	5948
Escherichia coli TOP 10F'	[pMS470 3_N12]	pMS470 3_N12.gb	3048
Escherichia coli TOP 10F'	[pMS470 10_K5]	pMS470 10_K5.gb	3049
Escherichia coli TOP 10F'	[pMS470 3_L6]	pMS470 3_L6.gb	3050
Escherichia coli TOP 10F'	[pMS470 3_L6/10_K5]	pMS470 3_L6 10_K5.gb	3051
Escherichia coli TOP 10F'	[pMS470 4_C2]	pMS470 4_C2.gb	3052
Escherichia coli TOP 10F'	[pMS470 2_O17]	pMS470 2_O17.gb	3053
Escherichia coli TOP 10F'	[pMS470 2_O7 SSMB]	pMS470 2_O7SSMB.gb	3054

[a] Designation in the dissertation [b] Designation in the lab journal

Vector pMS470 ER26syn (culture collection designation) or pMS470 P1syn (designation in the dissertation)



3.2 DNA and protein sequences

Part 1: Transposon Integration mediated Mutagenesis (TIM)

>PM1,17

ATGAACCATCCCGATATCGACACTCATTCACGAAATGCCGCCGCGCCTTTACCCTTTGTGTTGGTCCA
TGGAGCCTGGCATGGAGCCTGGGCCTATGAACGATTAGGAGCGGCGTTGGCGGCGCGTGGACATGCCA
GTGTCGCGCATGATTTACCCGCGCATGGAATTAATGCCCGATATCCGGCCGCGTTTTGGCAAGGAGAT
GCGCAAGCGTTAGCGCAAGAACCGTCTCCGGTCGCGGCCACAACTTTAGACGATTATACAGGACAAGT
GTTACGAGCGATCGATGCTGCTTGTGCTCTTGGTCACCCGAGAGTAGTGCTTGTAGGTCATAGTATGG
GTGGTGTAGCTATCACAGCCGCGGCTGAAAGAGCTCCGGAAAGAATCGCTGCTCTTGTATATCTTGCT
GCGTTCATGCCTGCTAGTGGTGTACCTGGTCTTGATTACGTGAGAGCTCCTGAAAACCATGGCGAAAT
GCTGGCCTCGCTGATCTGCGCCAGCCCTCGCGCGATCGGCGCGCTGCGCATCAACCCGGCCAGCCGCG
ACGCGGCCTATCTCGCCACGCTGAAGCAGGCGCTGTTCGAGGATGTTGACGAGGCGACGTTCCGCGCC
GTGACACGGCTGATGTCCTCGGACGTGCCGACCGCGCCATTCGCCACGCCGATCGCGACCACGGCCGA
GCGCTGGGGCTCGATCGCGCGCCACTACGTGACCTGCGCCGAGGATCGCGTGATCCTGCCGGCGCTGC
AGCGGCGCTTCATCGCCGAGGCCGACGCCTTCCTGCCCGAGCGGCCGACGCGCGTCCACGCACTCGAC

AGCAGCCATTCGCCGTTCCTGTCCCAGCCCGACACGCTCGCCGAGTTGCTGACGGGCATCGCGCGCAA
CACGGCGATCTGA

>PM1,17
MNHPDIDTHSRNAAAPLPFVLVHGAWHGAWAYERLGAALAARGHASVAHDLPAHGINARY
PAAFWQGDAQALAQEPSPVAATTLDDYTGQVLRAIDAACALGHPRVVLVGHSMGGVAITA
AAERAPERIAALVYLAAFMPASGVPGLDYVRAPENHGEMLASLICASPRAIGALRINPAS
RDAAYLATLKQALFEDVDEATFRAVTRLMSSDVPTAPFATPIATTAERWGSIARHYVTCA
EDRVILPALQRRFIAEADAFLPERPTRVHALDSSHSPFLSQPDTLAELLTGIARNTAI*

Part 3: 2-alkenal reductase

>P1syn
ATGACTGCGACCAACAAGCAGGTGATCCTGAAAGATTATGTATCTGGCTTCCCTACCGAATCTGATTT
CGATTTTACCACCACTACGGTAGAACTGCGCGTGCCGGAAGGCACTAACTCCGTTCTGGTGAAGAACC
TGTACCTGTCTTGTGATCCGTATATGCGTATCCGTATGGGCAAGCCGGACCCGTCTACCGCAGCGCTG
GCGCAGGCTTATACTCCGGGTCAGCCGATTCAGGGCTACGGTGTATCTCGTATCATCGAGTCTGGTCA
TCCGGATTATAAGAAAGGTGACCTGCTGTGGGGTATCGTGGCGTGGGAAGAGTACTCTGTTATCACCC
CGATGACCCACGCACATTTCAAAATCCAGCATACTGATGTACCGCTGTCTTATTACACCGGTCTGCTG
GGTATGCCTGGCATGACCGCGTACGCTGGTTTCTACGAAGTATGCTCTCCGAAAGAGGGCGAAACTGT
CTATGTTTCTGCAGCATCCGGCGCGGTGGGCCAACTGGTAGGCCAACTGGCCAAAATGATGGGTTGTT
ACGTGGTTGGCTCTGCGGGTAGCAAAGAGAAAGTTGACCTGCTGAAAACTAAATTCGGTTTCGATGAT
GCTTTCAATTATAAGGAAGAGAGCGATCTGACTGCGGCCCTGAAACGTTGTTTCCCGAACGGCATCGA
CATCTACTTCGAAAATGTGGGTGGCAAGATGCTGGATGCTGTGCTGGTTAACATGAATATGCACGGTC
GTATCGCGGTCTGCGGTATGATCTCTCAGTATAACCTGGAGAACCAGGAAGGTGTACATAACCTGTCC
AATATCATCTACAAACGCATCCGTATCCAGGGCTTCGTAGTCTCCGACTTCTACGATAAATATTCTAA
GTTCCTGGAGTTCGTGCTGCCGCACATCCGTGAAGGCAAGATCACGTATGTAGAGGACGTAGCGGACG
GCCTGGAGAAAGCTCCAGAGGCCCTGGTTGGTCTGTTCCATGGCAAGAATGTGGGCAAACAAGTCGTA
GTTGTCGCGCGTGAATAA

>P1syn
MTATNKQVILKDYVSGFPTESDFDFTTTTVELRVPEGTNSVLVKNLYLSCDPYMRIRMGK
PDPSTAALAQAYTPGQPIQGYGVSRIIESGHPDYKKGDLLWGIVAWEEYSVITPMTHAHF
KIQHTDVPLSYYTGLLGMPGMTAYAGFYEVCSPKEGETVYVSAASGAVGQLVGQLAKMMG
CYVVGSAGSKEKVDLLKTKFGFDDAFNYKEESDLTAALKRCFPNGIDIYFENVGGKMLDA
VLVNMNMHGRIAVCGMISQYNLENQEGVHNLSNIIYKRIRIQGFVVSDFYDKYSKFLEFV
LPHIREGKITYVEDVADGLEKAPEALVGLFHGKNVGKQVVVVARE*

>17_P15

ATGACTGCGACCAACAAGCAGGTGATCCTGAAAGATTATGTATCTGGCTTCCCTACCGAATCTGATTT
CGATTTTACCACCACTACGGTAGAACTGCGCGTGCCGGAAGGCACTAACTCCGTTCTGGTGAAGAACC
TGTACCTGTCTTGTGATCCGTATATGCGTATTTGTATGGGCAAGCCGGACCCGTCTACCGCAGCGCTG
GCGCAGGCTTATACTCCGGGTCAGCCGATTCAGGGCTACGGTGTATCTCGTATCATCGAGTCTGGTCA
TCCGGATTATAAGAAAGGTGACCTGCTGTGGGGTATCGTGGCGTGGGAAGAGTACTCTGTTATCACCC
CGAAGACCCACGCACATTACAAAATCCAGCATACTGATGTACCGCTGTCTTATTACACCGGTCTGCTG
GGTATGCCTGGCATGACCGCGTACGCTGGTTTCTACGAAGTATGCTCTCCGAAAGAGGGCGAAACTGT
CTATGTTTCTGCAGCATCCGGCGCGGTGGGCCAACTGGTAGGCCAACTGGCCAAAATGATGGGTTGTT
ACGTGGTTGGCTCTGCGGGTAGCAAAGAGAAAGTTGACCTGCTGAAAACTAAATTCGGTTTCGATGAT
GCTTTCAATTATAAGGAAGAGAGCGATCTGACTGCGGCCCTGAAACGTTGTTTCCCGAACGGCATCGA
CATCTACTTCGAAAATGTGGGTGGCAAGATGCTGGATGCTGTGCTGGTTAACATGAATATGCACGGTC
GTATCGCGGTCTGCGGTATGATCTCTCAGTATAACCTGGAGAACCAGGAAGGTGTACATAACCTGTCC
AATATCATCTACAAACGCATCCGTATCCAGGGCTTCGTAGTCTCCGACTTCTACGATAAATATTCTAA
GTTCCTGGAGTTCGTGCTGCCGCACATCCGTGAAGGCAAGATCACGTATGTAGAGGACGTAGCGGACG
GCCTGGAGAAAGCTCCAGAGGCCCTGGTTGGTCTGTTCCATGGCAAGAATGTGGGCAAACAAGTCGTA
GTTGTCGCGCGTGAATAA

>17_P15

MTATNKQVILKDYVSGFPTESDFDFTTTTVELRVPEGTNSVLVKNLYLSCDPYMRICMGK
PDPSTAALAQAYTPGQPIQGYGVSRIIESGHPDYKKGDLLWGIVAWEEYSVITPKTHAHY
KIQHTDVPLSYYTGLLGMPGMTAYAGFYEVCSPKEGETVYVSAASGAVGQLVGQLAKMMG
CYVVGSAGSKEKVDLLKTKFGFDDAFNYKEESDLTAALKRCFPNGIDIYFENVGGKMLDA
VLVNMNMHGRIAVCGMISQYNLENQEGVHNLSNIIYKRIRIQGFVVSDFYDKYSKFLEFV
LPHIREGKITYVEDVADGLEKAPEALVGLFHGKNVGKQVVVVARE*

>3_N12

ATGACTGCGACCAACAAGCAGGTGATCCTGAAAGATTATGTATCTGGCTTCCCTACCGAATCTGATTT
CGATTTTACCACCACTACGGTAGAACTGCGCGTGCCGGAAGGCACTAACTCCGTTCTGGTGAAGAACC
TGTACCTGTCTTGTGATCCGTATATGCGTATTTGTATGGGCAAGCCGGACCCGTCTACCGCAGCGCTG
GCGCAGGCTTATACTCCGGGTCAGCCGATTCAGGGCTACGGTGTATCTCGTATCATCGAGTCTGGTCA
TCCGGATTATAAGAAAGGTGACCTGCTGTGGGGTATCGTGGCGTGGGAAGAGTACTCTGTTATCACCC
CGAAGACCCACGCACATTACAAAATCCAGCATACTGATGTACCGCTGTCTTATTACACCGGTCTGCTG
GGTATGCCTGGCATGACCGCGTACGCTGGTTTCTACGAAGTATGCTCTCCGAAAGAGGGCGAAACTGT
CTATGTTTCTGCAGCATCCGGCGCGGTGGGCCAACTGGTAGGCCAACTGGCCAAAATGATGGGTTGTT
ACGTGGTTGGCTCTGCGGGTAGCAAAGAGAAAGTTGACCTGCTGAAAACTAAATTCGGTTTCGATGAT
GCTTTCAATTATAAGGAAGAGAGCGATCTGACTGCGGCCCTGAAACGTTGTTTCCCGAACGGCATCGA
CATCTACTTCGAAAATGTGGGTGGCAAGATGCTGGATGCTGTGCTGGTTAACATGAATATACACGGTC

GTATCGCGGTCTGCGGTATGATCTCTCAGTATAACCTGGAGAACCAGGAAGGTGTACATAACCTGTCC
AATATCATCTACAAACGCATCCGTATCCAGGGCTTCGTAGTCTCCGACTTCTACGATAAATATTCTAA
GTTCCTGGAGTTCGTGCTGCCGCACATCCGTGAAGGCAAGATCACGTATGTAGAGGACGTAGCGGACG
GCCTGGAGAAAGCTCCAGAGGCCCTGGTTGGTCTGTTCCATGGCAAGAATGTGGGCAAACAAGTCGTA
GTTGTCGCGCGTGAATAA

>3_12
MTATNKQVILKDYVSGFPTESDFDFTTTTVELRVPEGTNSVLVKNLYLSCDPYMRICMGK
PDPSTAALAQAYTPGQPIQGYGVSRIIESGHPDYKKGDLLWGIVAWEEYSVITPKTHAHY
KIQHTDVPLSYYTGLLGMPGMTAYAGFYEVCSPKEGETVYVSAASGAVGQLVGQLAKMMG
CYVVGSAGSKEKVDLLKTKFGFDDAFNYKEESDLTAALKRCFPNGIDIYFENVGGKMLDA
VLVNMNIHGRIAVCGMISQYNLENQEGVHNLSNIIYKRIRIQGFVVSDFYDKYSKFLEFV
LPHIREGKITYVEDVADGLEKAPEALVGLFHGKNVGKQVVVVARE*

>10_K5
ATGACTGCGACCAACAAGCAGGTGATCCTGAAAGATTATGTATCTGGCTTCCCTACCGAATCTGATTT
CGATTTTACCACCACTACGGTAGAACTGCGCGTGCCGGAAGGCACTAACTCCGTTCTGGTGAAGAACC
TGTACCTGTCTTGTGATCCGTATATGCGTATTTGTATGGGCAAGCCGGACCCGTCTACCGCAGCGCTG
GCGCAGGCTTATACTCCGGGTCAGCCGATTCAGGGCTACGGTGTATCTCGTATCATCGAGTCTGGTCA
TCCGGATTATAAGAAAGGTGACCTGCTGTGGGGTATCGTGGCGTGGGAAGAGTACTCTGTTATCACCC
CGAAGACCCACGCACATTACAAAATCCAGCATACTGATGTACCGCTGTCTTATTACACCGGTCTGCTG
GGTATGCCTGGCATGACCGCGTACGCTGGTTTCTACGAAGTATGCTCTCCGAAAGAGGGCGAAACTGT
CTATGTTTCTGCAGCATCCGGCGCGGTGGGCCAACTGGTAGGCCAACTGGCCAAAATGATGGGTTGTT
ACGTGGTTGGCTCTGCGGGTAGCAAAGAGAAAGTTGACCTGCTGAAAACTAAATTCGGTTTCGATGAT
GCTTTCAATTATAAGGAAGAGAGCGATCTGACTGCGGCCCTGAAACGTTGTTTCCCGAACGGCATCGA
CATCTACTTCGAAAATGTGGGTGGCAAGATGCTGGATGCTGTGCTGGTTAACATGAATGTGCACGGTC
GTATCGCGGTCTGCGGTATGATCTCTCAGTATAACCTGGAGAACCAGGAAGGTGTACATAACCTGTCC
AATATCATCTACAAACGCATCCGTATCCAGGGCTTCGTAGTCTCCGACTTCTACGATAAATATTCTAA
GTTCCTGGAGTTCGTGCTGCCGCACATCCGTGAAGGCAAGATCACGTATGTAGAGGACGTAGCGGACG
GCCTGGAGAAAGCTCCAGAGGCCCTGGTTGGTCTGTTCCATGGCAAGAATGTGGGCAAACAAGTCGTA
GTTGTCGCGCGTGAATAA

>10_K5
MTATNKQVILKDYVSGFPTESDFDFTTTTVELRVPEGTNSVLVKNLYLSCDPYMRICMGK
PDPSTAALAQAYTPGQPIQGYGVSRIIESGHPDYKKGDLLWGIVAWEEYSVITPKTHAHY
KIQHTDVPLSYYTGLLGMPGMTAYAGFYEVCSPKEGETVYVSAASGAVGQLVGQLAKMMG
CYVVGSAGSKEKVDLLKTKFGFDDAFNYKEESDLTAALKRCFPNGIDIYFENVGGKMLDA
VLVNMNVHGRIAVCGMISQYNLENQEGVHNLSNIIYKRIRIQGFVVSDFYDKYSKFLEFV
LPHIREGKITYVEDVADGLEKAPEALVGLFHGKNVGKQVVVVARE*

>3_L6

ATGACTGCGACCAACAAGCAGGTGATCCTGAAAGATTATGTATCTGGCTTCCCTACCGAATCTGATTT
CGATTTTACCACCACTACGGTAGAACTGCGCGTGCCGGAAGGCACTAACTCCGTTCTGGTGAAGAACC
TGTACCTGTCTTGTGATCCGTATATGCGTATTTGTATGGGCAAGCCGGACCCGTCTACCGCAGCGCTG
GCGCAGGCTTATACTCCGGGTCAGCCGATTCAGGGCTACGGTGTATCTCGTATCATCGAGTCTGGTCA
TCCGGATTATAAGAAAGGTGACCTGCTGTGGGGTATCGTGGCGTGGGAAGAGTACTCTGTTATCACCC
CGAAGACCCACGCACATTACAAAATCCAGCATACTGATGTACCGCTGTCTTATTACACCGGTCTGCTG
GGTATGCCTGGCATGACCGCGTACGCTGGTTTCTACGAAGTATGCTCTCCGAAAGAGGGCGAAACTGT
CTATGTTTCTGCAGCATCCGGCGCGGTGGGCCAACTGGTAGGCCAACTGGCCAAAATGATGGGTTGTT
ACGTGGTTGGCTCTGCGGGTAGCAAAGAGAAAGTTGACCTGCTGAAAACTAAATTCGGTTTCGATGAT
GCTTTCAATTATAAGGAAGAGAGCGATCTGACTGCGGCCCTGAAACGTTGTTTCCCGAACGGCATCGA
CATCTACTTCGAAAATGTGGGTGGCAAGATGCTGGATGCTGTGCTGGTTAACATGAATATGCACGGTC
GTATCGCGGTCTGCGGTATGGTCTCTCAGTATAACCTGGAGAACCAGGAAGGTGTACATAACCTGTCC
AATATCATCTACAAACGCATCCGTATCCAGGGCTTCGTAGTCTCCGACTTCTACGATAAATATTCTAA
GTTCCTGGAGTTCGTGCTGCCGCACATCCGTGAAGGCAAGATCACGTATGTAGAGGACGTAGCGGACG
GCCTGGAGAAAGCTCCAGAGGCCCTGGTTGGTCTGTTCCATGGCAAGAATGTGGGCAAACAAGTCGTA
GTTGTCGCGCGTGAATAA

>3_L6

MTATNKQVILKDYVSGFPTESDFDFTTTTVELRVPEGTNSVLVKNLYLSCDPYMRICMGK
PDPSTAALAQAYTPGQPIQGYGVSRIIESGHPDYKKGDLLWGIVAWEEYSVITPKTHAHY
KIQHTDVPLSYYTGLLGMPGMTAYAGFYEVCSPKEGETVYVSAASGAVGQLVGQLAKMMG
CYVVGSAGSKEKVDLLKTKFGFDDAFNYKEESDLTAALKRCFPNGIDIYFENVGGKMLDA
VLVNMNMHGRIAVCGMVSQYNLENQEGVHNLSNIIYKRIRIQGFVVSDFYDKYSKFLEFV
LPHIREGKITYVEDVADGLEKAPEALVGLFHGKNVGKQVVVVARE*

>3_L6/10_K5

ATGACTGCGACCAACAAGCAGGTGATCCTGAAAGATTATGTATCTGGCTTCCCTACCGAATCTGATTT
CGATTTTACCACCACTACGGTAGAACTGCGCGTGCCGGAAGGCACTAACTCCGTTCTGGTGAAGAACC
TGTACCTGTCTTGTGATCCGTATATGCGTATTTGTATGGGCAAGCCGGACCCGTCTACCGCAGCGCTG
GCGCAGGCTTATACTCCGGGTCAGCCGATTCAGGGCTACGGTGTATCTCGTATCATCGAGTCTGGTCA
TCCGGATTATAAGAAAGGTGACCTGCTGTGGGGTATCGTGGCGTGGGAAGAGTACTCTGTTATCACCC
CGAAGACCCACGCACATTACAAAATCCAGCATACTGATGTACCGCTGTCTTATTACACCGGTCTGCTG
GGTATGCCTGGCATGACCGCGTACGCTGGTTTCTACGAAGTATGCTCTCCGAAAGAGGGCGAAACTGT
CTATGTTTCTGCAGCATCCGGCGCGGTGGGCCAACTGGTAGGCCAACTGGCCAAAATGATGGGTTGTT
ACGTGGTTGGCTCTGCGGGTAGCAAAGAGAAAGTTGACCTGCTGAAAACTAAATTCGGTTTCGATGAT
GCTTTCAATTATAAGGAAGAGAGCGATCTGACTGCGGCCCTGAAACGTTGTTTCCCGAACGGCATCGA
CATCTACTTCGAAAATGTGGGTGGCAAGATGCTGGATGCTGTGCTGGTTAACATGAATGTGCACGGTC
GTATCGCGGTCTGCGGTATGGTCTCTCAGTATAACCTGGAGAACCAGGAAGGTGTACATAACCTGTCC

AATATCATCTACAAACGCATCCGTATCCAGGGCTTCGTAGTCTCCGACTTCTACGATAAATATTCTAA
GTTCCTGGAGTTCGTGCTGCCGCACATCCGTGAAGGCAAGATCACGTATGTAGAGGACGTAGCGGACG
GCCTGGAGAAAGCTCCAGAGGCCCTGGTTGGTCTGTTCCATGGCAAGAATGTGGGCAAACAAGTCGTA
GTTGTCGCGCGTGAATAA

>3_L6/10_K5
MTATNKQVILKDYVSGFPTESDFDFTTTTVELRVPEGTNSVLVKNLYLSCDPYMRICMGK
PDPSTAALAQAYTPGQPIQGYGVSRIIESGHPDYKKGDLLWGIVAWEEYSVITPKTHAHY
KIQHTDVPLSYYTGLLGMPGMTAYAGFYEVCSPKEGETVYVSAASGAVGQLVGQLAKMMG
CYVVGSAGSKEKVDLLKTKFGFDDAFNYKEESDLTAALKRCFPNGIDIYFENVGGKMLDA
VLVNMNVHGRIAVCGMVSQYNLENQEGVHNLSNIIYKRIRIQGFVVSDFYDKYSKFLEFV
LPHIREGKITYVEDVADGLEKAPEALVGLFHGKNVGKQVVVVARE*

>4_C2
ATGACTGCGACCAACAAGCAGGTGATCCTGAAAGATTATGTATCTGGCTTCCCTACCGAATCTGATTT
CGATTTTACCACCACTACGGTAGAACTGCGCGTGCCGGAAGGCACTAACTCCGTTCTGGTGAAGAACC
TGTACCTGTCTTGTGATCCGTATATGCGTATTTGTATGGGCAAGCCGGACCCGTCTACCGCAGCGCTG
GCGCAGGCTTATACTCCGGGTCAGCCGATTCAGGGCTACGGTGTATCTCGTATCATCGAGTCTGGTCA
TCCGGATTATAAGAAAGGTGACCTGCTGTGGGGTATCGTGGCGTGGGAAGAGTACTCTGTTATCACCC
CGAAGACCCACGCACATTACAAAATCCAGCATACTGATGTACCGCTGTCTTATTACACCGGTCTGCTG
GGTATGCCTGGCATGACCGCGTACGCTGGTTTCTACGAAGTATGCTCTCCGAAAGAGGGCGAAACTGT
CTATGTTTCTGCAGCATCCGGCGCGGTGGGCCAACTGGTAGGCCAACTGGCCAAAATGATGGGTTGTT
ACGTGGTTGGCTCTGCGGGTAGCAAAGAGAAAGTTGACCTGCTGAAAACTAAATTCGGTTTCGATGAT
GCTTTCAATTATAAGGAAGAGAGCGATCTGACTGCGGCCCTGAAACGTTGTTTCCCGAACGGCATCGA
CATCTACTTCGAAAATGTGGGTGGCAAGATGCTGGATGCTGTGCTGGTTAACATGAATATCCACGGTC
GTATCGCGGTCTGCGGTATGGTTTCTCAGTATAACCTGGAGAACCAGGAAGGTGTACATAACCTGTCC
AATATCATCTACAAACGCATCCGTATCCAGGGCTTCGTAGTCTCCGACTTCTACGATAAATATTCTAA
GTTCCTGGAGTTCGTGCTGCCGCACATCCGTGAAGGCAAGATCACGTATGTAGAGGACGTAGCGGACG
GCCTGGAGAAAGCTCCAGAGGCCCTGGTTGGTCTGTTCCATGGCAAGAATGTGGGCAAACAAGTCGTA
GTTGTCGCGCGTGAATAA

>4_C2
MTATNKQVILKDYVSGFPTESDFDFTTTTVELRVPEGTNSVLVKNLYLSCDPYMRICMGK
PDPSTAALAQAYTPGQPIQGYGVSRIIESGHPDYKKGDLLWGIVAWEEYSVITPKTHAHY
KIQHTDVPLSYYTGLLGMPGMTAYAGFYEVCSPKEGETVYVSAASGAVGQLVGQLAKMMG
CYVVGSAGSKEKVDLLKTKFGFDDAFNYKEESDLTAALKRCFPNGIDIYFENVGGKMLDA
VLVNMNIHGRIAVCGMVSQYNLENQEGVHNLSNIIYKRIRIQGFVVSDFYDKYSKFLEFV
LPHIREGKITYVEDVADGLEKAPEALVGLFHGKNVGKQVVVVARE*

>2_O17

ATGACTGCGACCAACAAGCAGGTGATCCTGAAAGATTATGTATCTGGCTTCCCTACCGAATCTGATTT
CGATTTTACCACCACTACGGTAGAACTGCGCGTGCCGGAAGGCACTAACTCCGTTCTGGTGAAGAACC
TGTACCTGTCTTGTGATCCGTATATGCGTATTTGTATGGGCAAGCCGGACCCGTCTACCGCAGCGCTG
GCGCAGGCTTATACTCCGGGTCAGCCGATTCAGGGCTACGGTGTATCTCGTATCATCGAGTCTGGTCA
TCCGGATTATAAGAAAGGTGACCTGCTGTGGGGTATCGTGGCGTGGGAAGAGTACTCTGTTATCACCC
CGAAGACCCACGCACATTACAAAATCCAGCATACTGATGTACCGCTGTCTTATTACACCGGTCTGCTG
GGTATGCCTGGCATGACCGCGTACGCTGGTTTCTACGAAGTATGCTCTCCGAAAGAGGGCGAAACTGT
CTATGTTTCTGCAGCATCCGGCGCGGTGGGCCAACTGGTAGGCCAACTGGCCAAAATGATGGGTTGTT
ACGTGGTTGGCTCTGCGGGTAGCAAAGAGAAAGTTGACCTGCTGAAAACTAAATTCGGTTTCGATGAT
GCTTTCAATTATAAGGAAGAGAGCGATCTGACTGCGGCCCTGAAACGTTGTTTCCCGAACGGCATCGA
CATCTACTTCGAAAATGTGGGTGGCAAGATGCTGGATGCTGTGCTGGTTAACATGAATACCCACGGTC
GTATCGCGGTCTGCGGTATGGTTTCTCAGTATAACCTGGAGAACCAGGAAGGTGTACATAACCTGTCC
AATATCATCTACAAACGCATCCGTATCCAGGGCTTCGTAGTCTCCGACTTCTACGATAAATATTCTAA
GTTCCTGGAGTTCGTGCTGCCGCACATCCGTGAAGGCAAGATCACGTATGTAGAGGACGTAGCGGACG
GCCTGGAGAAAGCTCCAGAGGCCCTGGTTGGTCTGTTCCATGGCAAGAATGTGGGCAAACAAGTCGTA
GTTGTCGCGCGTGAATAA

>2_O17

MTATNKQVILKDYVSGFPTESDFDFTTTTVELRVPEGTNSVLVKNLYLSCDPYMRICMGK
PDPSTAALAQAYTPGQPIQGYGVSRIIESGHPDYKKGDLLWGIVAWEEYSVITPKTHAHY
KIQHTDVPLSYYTGLLGMPGMTAYAGFYEVCSPKEGETVYVSAASGAVGQLVGQLAKMMG
CYVVGSAGSKEKVDLLKTKFGFDDAFNYKEESDLTAALKRCFPNGIDIYFENVGGKMLDA
VLVNMNTHGRIAVCGMVSQYNLENQEGVHNLSNIIYKRIRIQGFVVSDFYDKYSKFLEFV
LPHIREGKITYVEDVADGLEKAPEALVGLFHGKNVGKQVVVVARE*

>2_O7SSMB

ATGACTGCGACCAACAAGCAGGTGATCCTGAAAGATTATGTATCTGGCTTCCCTACCGAATCTGATTT
CGATTTTACCACCACTACGGTAGAACTGCGCGTGCCGGAAGGCACTAACTCCGTTCTGGTGAAGAACC
TGTACCTGTCTTGTGATCCGTATATGCGTATTTGTATGGGCAAGCCGGACCCGTCTACCGCAGCGCTG
GCGCAGGCTTATACTCCGGGTCAGCCGATTCAGGGCTGGGGTGTATCTCGTATCATCGAGTCTGGTCA
TCCGGATTATAAGAAAGGTGACCTGCTGTGGGGTATCGTGGCGTGGGAAGAGTACTCTGTTATCACCC
CGAAGACCCACGCACATTACAAAATCCAGCATACTGATGTACCGCTGTCTTATTACACCGGTCTGCTG
GGTATGCCTGGCATGACCGCGTACGCTGGTTTCTACGAAGTATGCTCTCCGAAAGAGGGCGAAACTGT
CTATGTTTCTGCAGCATCCGGCGCGGTGGGCCAACTGGTAGGCCAACTGGCCAAAATGATGGGTTGTT
ACGTGGTTGGCTCTGCGGGTAGCAAAGAGAAAGTTGACCTGCTGAAAACTAAATTCGGTTTCGATGAT
GCTTTCAATTATAAGGAAGAGAGCGATCTGACTGCGGCCCTGAAACGTTGTTTCCCGAACGGCATCGA
CATCTACTTCGAAAATGTGGGTGGCAAGATGCTGGATGCTGTGCTGGTTAACATGAATATGCACGGTC
GTATCGCGGTCTGCGGTATGATCTCTCAGTATAACCTGGAGAACCAGGAAGGTGTACATAACCTGTCC

AATATCATCTACAAACGCATCCGTATCCAGGGCTTCGTAGTCTCCGACTTCTACGATAAATATTCTAA
GTTCCTGGAGTTCGTGCTGCCGCACATCCGTGAAGGCAAGATCACGTATGTAGAGGACGTAGCGGACG
GCCTGGAGAAAGCTCCAGAGGCCCTGGTTGGTCTGTTCCATGGCAAGAATGTGGGCAAACAAGTCGTA
GTTGTCGCGCGTGAATAA

>2_O7SSMB
MTATNKQVILKDYVSGFPTESDFDFTTTTVELRVPEGTNSVLVKNLYLSCDPYMRICMGK
PDPSTAALAQAYTPGQPIQGWGVSRIIESGHPDYKKGDLLWGIVAWEEYSVITPKTHAHY
KIQHTDVPLSYYTGLLGMPGMTAYAGFYEVCSPKEGETVYVSAASGAVGQLVGQLAKMMG
CYVVGSAGSKEKVDLLKTKFGFDDAFNYKEESDLTAALKRCFPNGIDIYFENVGGKMLDA
VLVNMNMHGRIAVCGMISQYNLENQEGVHNLSNIIYKRIRIQGFVVSDFYDKYSKFLEFV
LPHIREGKITYVEDVADGLEKAPEALVGLFHGKNVGKQVVVVARE*

3.3 Primers

Part 1: Transposon Integration mediated Mutagenesis (TIM)

Number	Primer Name	Sequence in 5'-3' orientation	Use
P05-250	*Mu*ins1bpAarI	GCTCAGATCTGGCAGGTGGCACGAAAAACG	PCR
P05-356	*Mu*del1bpAarI	GGAGCTCAGATCTGCAGGTGCGCACGAAAAACG	PCR
P04-137	6xHisG	GGCCATCATCATCATCATCACGGC	linker
P04-172	6xHisGrev	GCCGTGATGATGATGATGATGGCC	linker
Nr. 0664	TacNEU	GAATTCGAGCTCGGTAC	insert sequencing
Nr. 0506	TacSTOP	CATCCGCCAAAACAGCC	insert sequencing
P06-513	histagTIMfw	GCCATCATCATCATCATCAC	sequencing

Part 3: 2-alkenal reductase

Number	Primer Name	Sequence in 5'-3' orientation	Use
P08-425	P1fw	ATGACTGCGACCAACAAGC	Mutazyme II PCR
P08-426	P1rv	TTATTCACGCGCGACAACTAC	Mutazyme II PCR
P08-255	P1*Nde*I[a)] ER26*Nde*I[b)]	CTAACATATGACTGCGACCAACAAGC	Taq PCR

P08-256	P1*Hin*dIII[a)] ER26*Hin*dIII[b)]	CTAGAAGCTTTTATTCACGCGCGAC	Taq PCR
P09-58	SSMB_Tyr81_fw*	CGATTCAGGGCNNKGGTGTATCTCGTATCATC	randomize: Y81
P09-59	SSMB_Tyr81_rv*	GATGATACGAG ATACACCMNNGCCCTGAATCG	randomize: Y81
P09-60	SSMB_Tyr290_fw*	CGTAGTCTCCGACTTCNNKGATAAATATTCTAA	randomize: Y290
P09-61	SSMB_Tyr290_rv*	CTTAGAATATTTATCMNNGAAGTCGGAGACTAC	randomize: Y290
P08-257	pMSfw	GTGAGCGGATAACAATTTCACACA	insert sequencing
P08-258	pMSrv	GTTTTATCAGACCGCTTCTGCG	insert sequencing

[a)] Designation in the dissertation [b)] Designation in the lab journal *SSM2 in the lab journal

3.4 Materials and equipment (additional information)

Part 1: Transposon Integration mediated Mutagenesis (TIM)

Medium	Ingredients
SOC	20g/L Bacto™ Tryptone, 0,58g/L NaCl, 5g/L Bacto™ Yeast Extract, 2g/L $MgCl_2$, 0.81g/L KCl, 2,46g/L $MgSO_4$
Low Salt Luria-Bertani (LB)*	10g/L Bacto™ Tryptone, 5g/L Bacto™ Yeast Extract, 5g/L NaCl
LB Agar plates*	10g/L Bacto™ Tryptone, 5g/L Bacto™ Yeast Extract, 5g/L NaCl, 15g/L agarose

* Antibiotics, if needed, were added below 60°C (final concentration ampicillin: 100 mg/L and kanamycin 10 mg/L, respectively)

Instrument	Producer
Gene Pulser and 0.2cm cuvettes	Bio-Rad Laboratories, Vienna, Austria
Electrophoresis unit	Biozyme ComPhor L Midi
Microwave oven	e.g. Samsung M945
Model 250 / 2.5 Power supply	BIORAD, CA, USA
Model 100 / 500 Power supply	BIORAD, CA, USA

Thermomixer comfort	Eppendorf AG, Hamburg, Germany
Titramax 1000	Heidolph Instruments GmbH, Schwabach, Germany

Part 2: Reductive amination

Material	Producer
TLC aluminium sheets, 20x20cm, silica gel 60 F^{254}	MERCK, Darmstadt, Germany

Part 3: 2-alkenal reductase

Material	Producer
LE Agarose	Biozym, Hessisch Oldendorf, Germany
electroporation cuvettes, 0.2cm, EP-102	Cell projects, Harrietsham, Kent, UK.
96 well, PS, F-bottom	Greiner Bio One (Kremsmünster, Austria)
96 well, PS, V-bottom	Greiner Bio One (Kremsmünster, Austria)
PS-microplate 384 well	Greiner Bio One (Kremsmünster, Austria)
UV-Star 96 well plate	Greiner Bio One (Kremsmünster, Austria)
Cellstar®serological pipettes 5 ml, 10 ml, 25 ml	Eppendorf AG, Hamburg, Germany

Instrument	Producer
MicroPulser™	Bio-Rad Laboratories, Vienna, Austria
Electrophoresis unit	Sub-Cell® GT, BIORAD, CA, USA
Microwave oven	e.g. Silva MW-2080
PowerPac Basic™ Power supply	BIORAD, CA, USA
GP 3202 Precision balance	Sartorius, Goettingen, Germany
Electronic balance ABS 220-4	Kern & Sohn GmbH Balingen-Frommern, Germany
pH-Meter inoLab® pH 720	WTW, Weilheim, Germany
Thermomixer comfort	Eppendorf AG, Hamburg, Germany
GeneAmp ® System 2400	Applied Biosystems, Foster City, CA, USA
GeneAmp ® System 2700	Applied Biosystems, Foster City, CA, USA
Centrifuge 5810R	Eppendorf AG, Hamburg, Germany
Centrifuge 5415R	Eppendorf AG, Hamburg, Germany
Biohit Proline ® single-channel electronic pipettor 0.2 – 10 µl	Biohit Plc., Helsinki, Finland
Biohit Proline ® single-channel electronic pipettor 0.2 – 10 µl	Biohit Plc., Helsinki, Finland
Biohit Proline ® multi-channel electronic pipettor, 8 channel 5-100 µl, 8 channel 50-1200 µl	Biohit Plc., Helsinki, Finland
Easypet® 4420	Eppendorf AG, Hamburg, Germany

3.5 Standard for protein gels

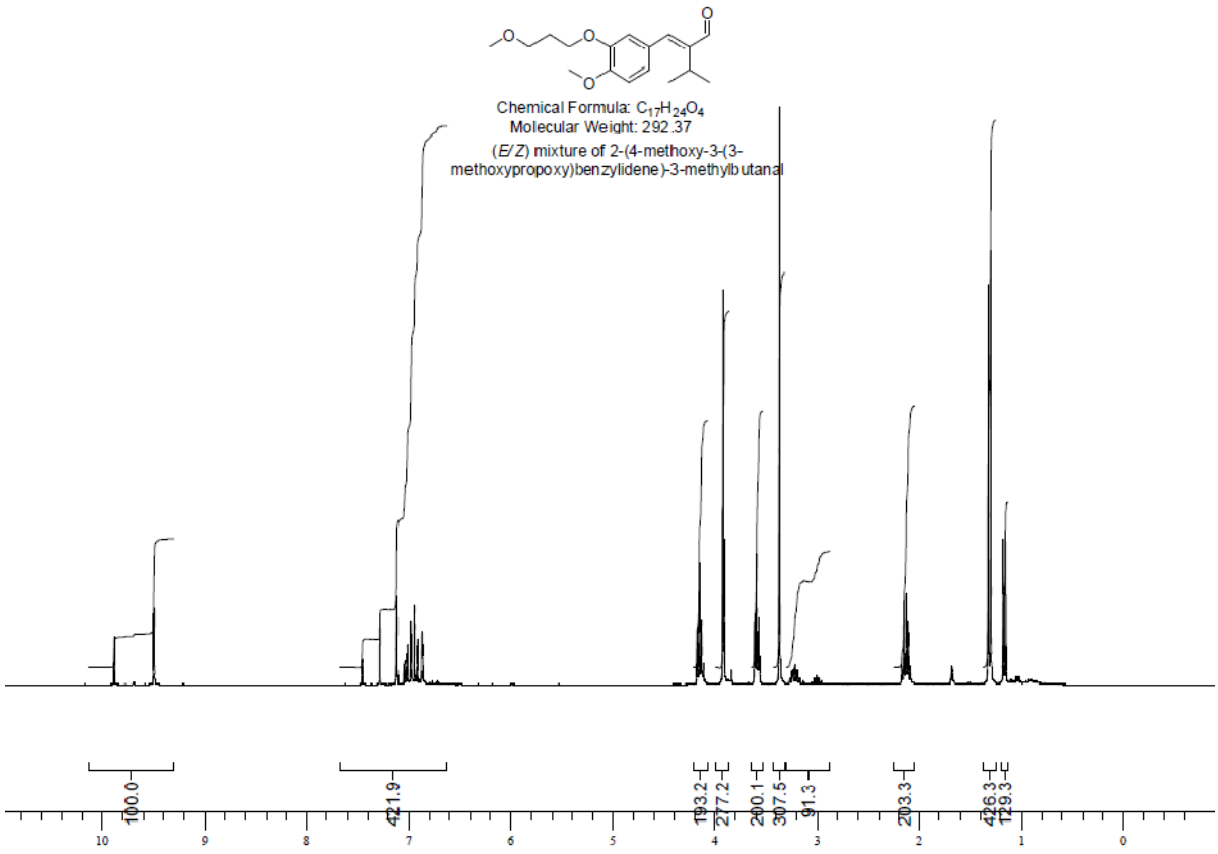
SeeBlue® Plus2 Pre-Stained Standard from Invitrogen (Catalog no. LC5925)

NuPAGE® Novex
Bis-Tris 4-12% Gel

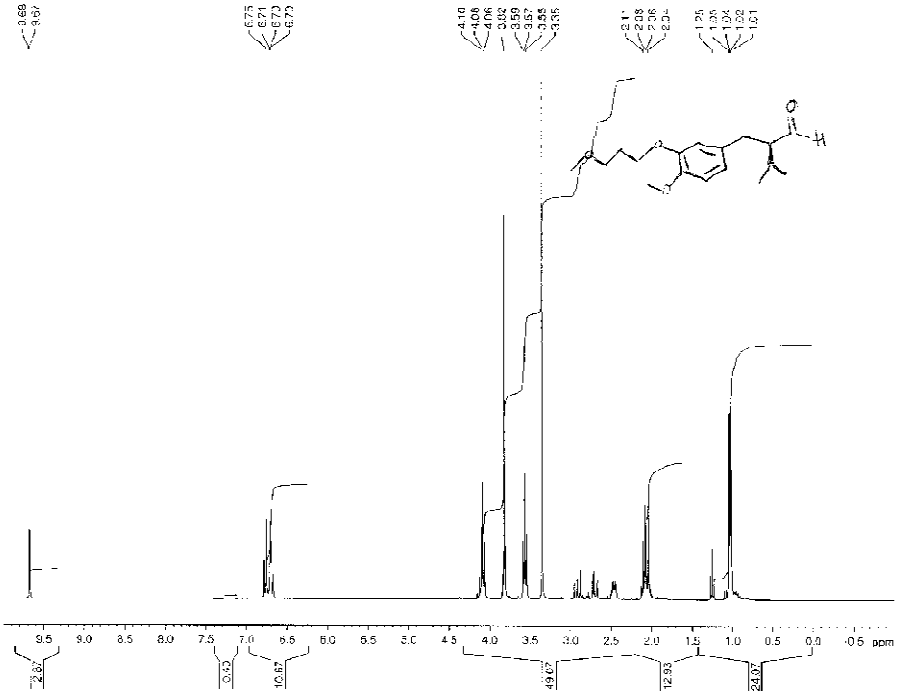
Protein	Approximate Molecular Weights (kDa)				
	Tris-Glycine	Tricine	NuPAGE® MES	NuPAGE® MOPS	NuPAGE® Tris-Acetate
Myosin	250	210	188	191	210
Phosphorylase	148	105	98	97	111
BSA	98	78	62	64	71
Glutamic Dehydrogenase	64	55	49	51	55
Alcohol Dehydrogenase	50	45	38	39	41
Carbonic Anhydrase	36	34	28	28	n/a
Myoglobin Red	22	17	17	19	n/a
Lysozyme	16	16	14	14	n/a
Aprotinin	6	7	6	n/a	n/a
Insulin, B Chain	4	4	3	n/a	n/a

3.6 NMR data of references – 2-alkenal reductase

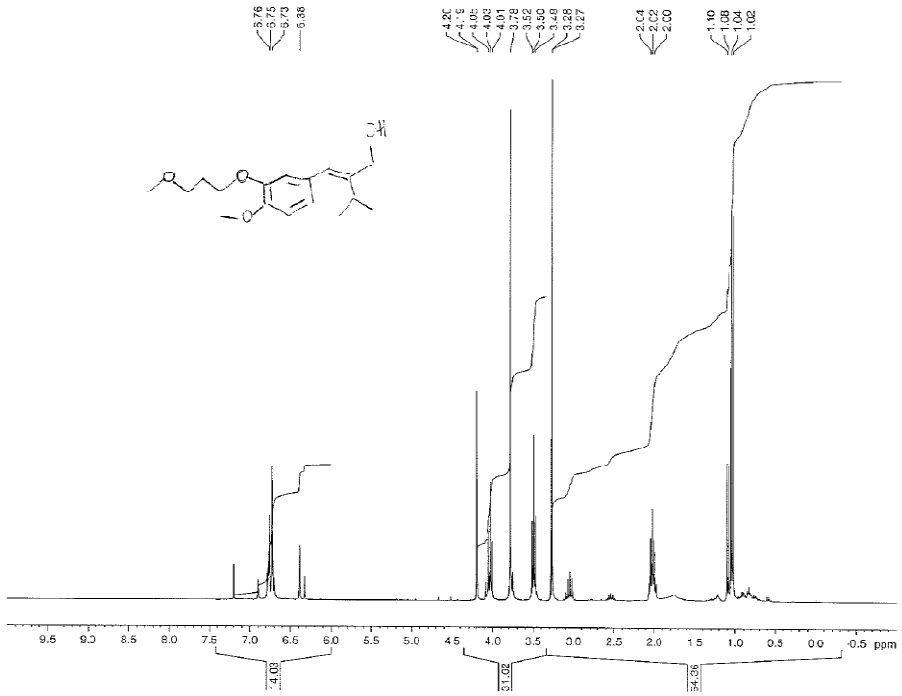
(E/Z) Mixture of 2-[4-methoxy-3-(3-methoxypropoxy) benzylidene]-3-methylbutanal (**1a**)



Saturated aldehyde product (**1b**)



E/Z allylalcohol (**1d**)



Additional Information on CD

Designation: Dissertation Brigitte Höller

Dissertation in word and PDF format

DNA and protein sequences (wild type and most important variants)

Presentations from TIM, reductive amination, 2-alkenal reductase and final presentation

Endnote library and PDF files (literature)

VNTI files

Project folders

3.7 Table of figures

CURRICULUM VITAE

Persönliche Daten:

Brigitte Mathilde Höller

Grazer Vorstadt 91, 8570 Voitsberg, Österreich

Tel.: +43 676/9742080

Email: brigitte.hoeller@gmx.at

Geburtsdatum: 25.03.1982

Geburtsort: 8570 Voitsberg

Staatsangehörigkeit: Österreich

Schulausbildung:

1988 - 1992 Volksschule Voitsberg

1992 - 2000 Bundesgymnasium und Bundesrealgymnasium Köflach

Akademische Ausbildung:

2000 - 2006 Diplomstudium der Biologie (Studienzweig Mikrobiologie) an der Karl-Franzens-Universität Graz

Diplomarbeit: **The impact of IncI1 plasmids on bacterial biofilm development and their prevalence in natural *E. coli* isolates**

Betreuerin: Ao. Univ.-Prof. Dr. rer. nat. Ellen L. Zechner

Karl-Franzens-Universität Graz, Institut für Mikrobiologie

2006 – Doktoratsstudium an der Technischen Universität Graz

Dissertation: **New Oxidoreductases for Biocatalysis**

Betreuer: Ao. Univ.-Prof. Mag. rer. nat. Dr. rer. nat. Anton Glieder

Technische Universität Graz, Institut für Molekulare Biotechnologie, Forschungszentrum Angewandte Biokatalyse Kompetenzzentrum GmbH

Berufserfahrungen und Kompetenzen:

Ferialpraktika: Stadtwerke Voitsberg, Kaufmännische Verwaltung

Institut für Hygiene Graz, Abteilung Wasserhygiene und Mikroökologie

Weitere Tätigkeiten: Tutorin am Institut für Molekulare Biotechnologie (SS07, WS07, WS08)

Betreuung von Sommerstudenten und einem Diplomanden

Projektleitungserfahrung (Angewandte Biokatalyse)

Erste Hilfe Beauftragte (Angewandte Biokatalyse)

Sonstige Fähigkeiten:

Sprachen:	Deutsch (Muttersprache), Englisch (sehr gute Kenntnisse in Wort und Schrift), Italienisch (Basis)
EDV:	MS office, Vector NTI, SeqMan, Pymol, ChemSketch, Endnote
Führerschein:	Klasse B

Teilnahme an Konferenzen:

2006: International Symposium on the Genetics of Industrial Microorganisms (GIM), Prag, Tschechien

Multi Step Enzyme Catalysed Processes (MECP), Graz, Austria

2007: Society for Industrial Microbiology, Annual Meeting and Exhibition 2007 (SIM), Denver, CO, USA

VAAM Annual Meeting, Osnabrück, Deutschland

2009: ProStab2009, 8th International Conference on Protein Stabilisation, Graz, Austria

---

Präsentationen:

Poster: New Methods for Directed Evolution: TIM (Transposon Integration Mediated Mutagenesis), GIM 2006

Random Insertion of Tags into Biocatalytic Enzymes, SIM 2007

Vortrag: Enzyme Development by Transposon-Based Random Insertions and Deletions, VAAM Annual Meeting 2007

Publikationen:

Reisner, A., Höller, B.M., Molin, S., and Zechner, E.L. **Synergistic effects in mixed E. coli biofilms: conjugative plasmid transfer drives biofilm expansion** (2006) Journal of Bacteriology, 188 (10), pp. 3582-3588.

Hoeller B.M., Reiter B., Abad S., Graze I., Glieder A. **Random tag insertions by Transposon Integration mediated Mutagenesis (TIM)** (2008), Journal of Microbiological Methods, 75 (2), pp.251-257.

Lavandera I., Höller B., Kern A., Ellmer U., Glieder A., de Wildeman S., Kroutil W. **Asymmetric anti-Prelog reduction of ketones catalysed by Paracoccus pantotrophus and Comamonas sp. cells via hydrogen transfer** (2008), Tetrahedron Asymmetry, 19 (16), pp.1954-1958.

Persönliche Interessen:

Sport, Musik, Theater, Bücher

Referenzen:

- Prof. Anton Glieder
  Institut für Molekulare Biotechnologie
  Technische Universität Graz
  8010 Graz, Österreich
  Petersgasse 14/2
  Tel: +43 316 873 4074
  a.glieder@tugraz.at

- Prof. Ellen L. Zechner
  Institut für Molekulare Biowissenschaften
  Karl-Franzens-Universität Graz
  Humboldtstrasse 50/1
  8010 Graz, Österreich
  Tel: +43 316 380 5624
  ellen.zechner@uni-graz.at

- Prof. Helmut Schwab
  Institut für Molekulare Biotechnologie
  Technische Universität Graz
  8010 Graz, Österreich
  Petersgasse 14/5
  Tel: +43 316 873 4070
  helmut.schwab@tugraz.at