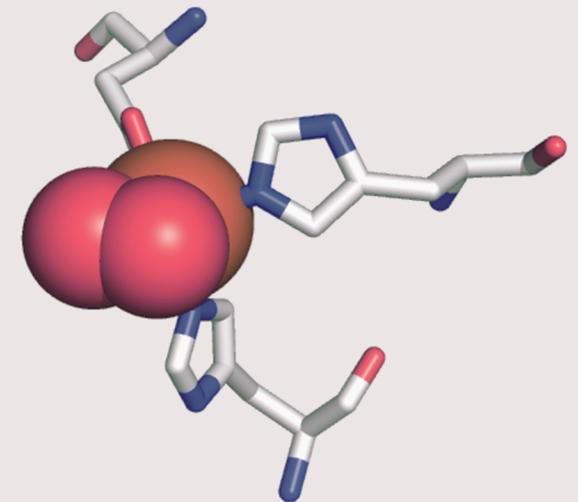
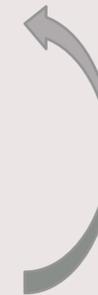
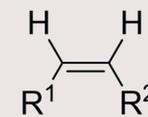
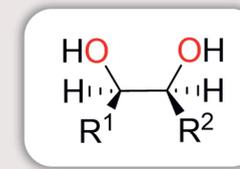
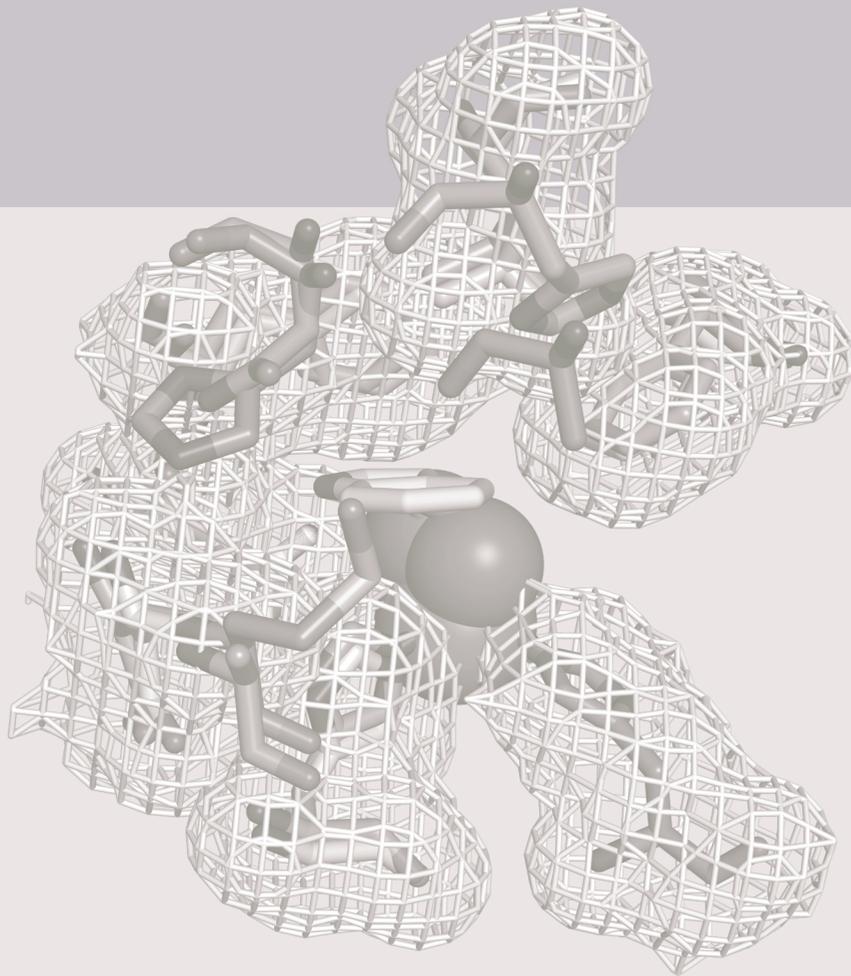


# ENZYMATIC ASYMMETRIC DIHYDROXYLATION OF ALKENES

ENZYMATISCHE ASYMMETRISCHE DIHYDROXYLIERUNG VON ALKENEN



CHRISTINE GALLY

Institut für Technische Biochemie - Universität Stuttgart

# **Enzymatic asymmetric dihydroxylation of alkenes**

Enzymatische asymmetrische Dihydroxylierung von Alkenen

Von der Fakultät 4: Energie-, Verfahrens- und Biotechnik der Universität Stuttgart  
zur Erlangung der Würde eines Doktors der Naturwissenschaften (Dr. rer. nat.)  
genehmigte Abhandlung.

Vorgelegt von  
**Christine Gally**  
aus  
Karlsruhe

Hauptberichter: Prof. Dr. Bernhard Hauer  
Mitberichter: Prof. Dr. Dieter Wolf  
Vorsitzender: Prof. Dr. Karl-Heinrich Engesser

Tag der Verteidigung: 31.05.2016

Institut für Technische Biochemie der Universität Stuttgart

**2016**



## COVER

### FRONT:

The catalytic non-heme iron (orange) with side-on bound dioxygen (red) coordinated by the 2-His-1-carboxylate facial triad in the active site of naphthalene dioxygenase from *Pseudomonas putida* NCIB 9816-4 (PDB code: 1O7N).

### BACK:

Alignment of amino acid residues within a distance of 8 Å from the non-heme iron of naphthalene dioxygenase from *Pseudomonas putida* NCIB 9816-4 (PDB code: 1O7N) and cumene dioxygenase from *Pseudomonas fluorescens* IP01 (PDB code: 1WQL).

The present thesis was prepared under the supervision of Prof. Dr. Bernhard Hauer in the period of January 2012 to June 2015 at the Institute of Technical Biochemistry at the University of Stuttgart, Germany.

Parts of this thesis are published in:

Gally, C., Nestl, B. M. & Hauer, B. Engineering Rieske Non-Heme Iron Oxygenases for the Asymmetric Dihydroxylation of Alkenes. *Angew. Chem. Int. Ed.* **54**, 12952–12956 (2015)

## DECLARATION OF AUTHORSHIP

I hereby declare that the present thesis entitled "Enzymatic asymmetric dihydroxylation of alkenes" is the result of my own work, that all sources used or quoted have been indicated, and that I have not used any illegitimate means. I further declare that I have not submitted this thesis for a degree in some form or another.

Ich versichere, dass ich die vorliegende Arbeit mit dem Titel „Enzymatische asymmetrische Dihydroxylierung von Alkenen“ selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Aus fremden Quellen entnommene Passagen und Gedanken sind als solche kenntlich gemacht. Des Weiteren bestätige ich ausdrücklich, dass die hier vorgelegte Dissertation nicht in gleicher oder ähnlicher Form bei einer anderen Institution zur Erlangung eines akademischen Grades eingereicht wurde.

Stuttgart, 15.01.2015

Christine Gally

## ACKNOWLEDGEMENTS

An dieser Stelle möchte ich mich gerne bei den Leuten bedanken, die die Anfertigung dieser Arbeit ermöglicht und mit dazu beigetragen haben, dass die Dissertation in ihrer jetzigen Form entstanden ist.

Mein großer Dank gilt dabei Herrn Prof. Dr. Bernhard Hauer, der es mir ermöglicht hat, meine Promotion am Institut für Technische Biochemie durchzuführen. Ich möchte mich herzlich für das spannende und faszinierende Themengebiet bedanken, auf dem ich meine Promotion durchführen konnte sowie für die Unterstützung während meiner Dissertation. Außerdem möchte ich für die Möglichkeit danken, dass ich an zahlreichen internationale Konferenzen und Projekttreffen teilnehmen konnte.

Bedanken möchte ich mich auch bei Herrn Prof. Dr. Dieter Wolf für die freundliche Übernahme des Zweitgutachtens meiner Arbeit und bei Herrn Prof. Dr. Karl-Heinrich Engesser für die freundliche Übernahme des Prüfungsvorsitzes.

Mein besonderer Dank gilt Dr. Bettina Nestl für die hervorragende Betreuung während meiner Promotion, die hilfreichen Diskussionen sowie für das kritische Korrekturlesen der vorliegenden Arbeit. Vielen Dank für die Unterstützung und das Vertrauen in schwierigen Phasen des Projektes.

Ebenso möchte ich mich bei Dr. Bernd Nebel für die hilfreichen Ratschläge in analytischen Fragestellungen sowie die Instandhaltung und Pflege der GCs, HPLCs und aller anderen analytischen Geräte bedanken.

Bei Sven Richter möchte ich mich für seine Hilfestellung bei der Fermentation sowie den vielen interessanten Diskussionen zu diesem Thema herzlich bedanken.

Bedanken möchte ich mich auch bei Dr. Janosch Klebensberger und Dr. Sandra Facey für Ihre Hilfe in molekular- und mikrobiologischen Fragen. Dr. Sandra Facey danke ich außerdem für das Bereitstellen von Vektoren und Stämmen sowie die jährliche Zusammenarbeit beim TBC I Praktikum.

Prof. Dr. Jürgen Pleiss und Dr. Constantin Vogel danke ich für die Kooperation im bioinformatischen Bereich, die zur Identifizierung des konservierten Motivs in der Dioxygenase-Datenbank geführt hat.

Des Weiteren danke ich Dr. Birgit Claasen für die Hilfestellungen beim Auswerten komplizierter NMR Spektren.

Ganz herzlich möchte ich mich außerdem bei allen jetzigen und ehemaligen Mitarbeitern des ITB für die freundliche und entspannte Atmosphäre bedanken. Vielen Dank Jenny, Sandra V., Thorsten, Andy, Chris (Schneck), Sandra N., Jule, Lisa, Sebastian, Leonie, Jan, Philipp T., Mihaela, Matthias, Silvi, Sara, Conny, Dominique, Martin, Maike, Rebecca, Max, Christian, Stephan, Konrad, Tobi, Björn, Mina, Melanie, Jens, Nico, Daniel, Marko, Wouter, Franz, Per-Olof, Julia und Sumire für die schöne Zeit während meiner Promotion. Besonderen Dank hierbei an meine Kollegen aus dem kleinen Labor, Silke, Jörg und Miri, die stets ein offenes Ohr für meine Sorgen und Probleme hatten, mich motiviert haben und mir immer zur Seite gestanden sind – Ohne euch wäre es nur halb so gut gewesen.

Bedanken möchte ich mich auch bei den Studenten Dominik Essig, Leonie Weinmann, Andreas Kunzendorfer und Thomas Weisenburger, die an verschiedenen Teilen dieses Projektes mitgearbeitet haben und von mir betreut wurden.

Der europäischen Union und unseren Kooperationspartner gilt mein Dank für die Finanzierung dieses Projektes im Rahmen von „CHEM21“. Des Weiteren möchte ich mich bei Prof. Dr. Rebecca Parales, Prof. Dr. David Leak und Prof. Dr. Hideaki Nojiri für die Bereitstellung der in dieser Arbeit verwendeten Plasmide bedanken.

Zum Schluss möchte ich mich noch bei meiner Familie und insbesondere bei Philipp für die Hilfe und Unterstützung während meines Studiums und meiner Promotion bedanken, die zu jeder Zeit für mich da waren.

## TABLE OF CONTENTS

DECLARATION OF AUTHORSHIP.....	5
ACKNOWLEDGEMENTS .....	6
TABLE OF CONTENTS .....	8
ABBREVIATIONS .....	10
ABSTRACT .....	11
ZUSAMMENFASSUNG.....	13
1. INTRODUCTION.....	15
1.1. Vicinal diols - Building blocks in the chemical and pharmaceutical industry.....	15
1.2. Synthesis of vicinal <i>cis</i> -diols.....	16
1.2.1. Chemical catalysts: Metal-catalyzed asymmetric dihydroxylation (AD) .....	16
1.2.2. Biocatalytic access to <i>cis</i> -1,2-diols .....	20
1.3. Rieske non-heme iron oxygenases (ROs) .....	23
1.3.1. Classification of RO systems.....	23
1.3.2. Substrate scope and reaction spectrum .....	25
1.3.3. Structure and catalytic mechanism of ROs .....	28
1.3.4. Applications of ROs .....	31
1.4. Protein engineering – Different strategies for the tailoring of ROs.....	32
1.4.1. Rational enzyme design .....	33
1.4.2. Directed evolution.....	33
1.4.3. Semi-rational design: Combining directed evolution and rational design.....	35
1.5. Aim of the work.....	37
2. EXPERIMENTAL SECTION .....	38
2.1. Materials .....	38
2.2. Plasmids, strains and storage of <i>E. coli</i> cultures.....	38
2.3. Methods .....	39
2.3.1. Restriction enzyme-based gene cloning - Reconstruction of plasmid pIP107D.....	39
2.3.2. Site-directed mutagenesis .....	40
2.3.3. Site-saturation mutagenesis and indole screening assay .....	41
2.3.4. Preparation of chemically competent <i>E. coli</i> cells and transformation of <i>E. coli</i> ....	42
2.3.5. Heterologous expression of RO wild type enzymes and variants .....	42
2.3.6. <i>In vivo</i> biotransformations with wild type and mutant enzymes.....	43
2.3.7. Up-scaling of ( <i>R</i> )-limonene hydroxylation to 4 L.....	43
2.3.8. Microbial toxicity of (1 <i>R</i> ,5 <i>S</i> )-carveol on <i>E. coli</i> JM109 .....	44
2.3.9. GC, HPLC and NMR analytics.....	44
2.3.10. Chemical synthesis of product standards.....	48
2.3.11. Identification of biotransformation products by NMR and GC-MS analysis .....	54
2.3.12. Docking .....	61

---

3. RESULTS .....	62
3.1. Oxyfunctionalization of alkenes by wild type ROs.....	62
3.1.1. Cloning and expression of RO genes.....	62
3.1.2. Pre-screening of different olefins .....	63
3.1.3. Characterization of ROs towards a panel of selected alkene substrates.....	65
3.2. Rational design – Optimization of ROs for selective olefin oxyfunctionalization .....	68
3.2.1. Selection of amino acid residues by phylogenetic comparison .....	68
3.2.2. Rational design - Characterization of generated RO variants .....	70
3.3. Site-saturation mutagenesis of CDO - Creation of a focused mutant library .....	74
3.3.1. Exploring the influence of various amino acid residues at position 232.....	74
3.3.2. Site-saturation mutagenesis at additional positions in CDO .....	79
3.4. Semi-preparative biotransformations with CDO M232A as catalyst .....	84
3.4.1. Biotransformations with resting cells in 70 mg scale .....	84
3.4.2. Biotransformation of ( <i>R</i> )-limonene with growing cells in a bioreactor (4 L).....	84
4. DISCUSSION.....	88
4.1. Oxyfunctionalization of alkene substrates by RO wild type enzymes .....	88
4.2. Influence of different active site residues on the RO-catalyzed alkene oxidation .....	92
4.3. Semi-preparative biotransformations with CDO M232A .....	97
4.4. RO-catalyzed alkene oxyfunctionalization <i>versus</i> Sharpless AD and Riley oxidation – Benefits and limitations.....	100
4.4.1. Sharpless AD .....	100
4.4.2. Riley oxidation .....	102
5. CONCLUSIONS AND OUTLOOK.....	104
6. REFERENCES.....	107
7. SUPPLEMENTARY .....	119
7.1. Plasmids .....	119
7.2. Sequencing primers.....	120
7.3. Genes .....	121
7.4. Sequence alignments .....	125
7.5. Stereoselectivities.....	129
7.6. Conservation level of positions F278, I288, I336 and F378 in CDO .....	145
7.7. NMR spectra.....	146
8. CURRICULUM VITAE .....	200

## ABBREVIATIONS

°C	degree Celsius	IS	internal standard
Å	Ångström	L	liter
AD	asymmetric dihydroxylation	LB	lysogeny broth
ADH	alcohol dehydrogenase	μ	micro
BAL	benzaldehyde lyase	m	mili
BDO	benzene dioxygenase	M	molar
bp	base pair	Mn	manganese
BphDO	biphenyl dioxygenase	nm	nanometer
CAST	combinatorial active-site saturation test	NAD(P)H	nicotinamide adenine dinucleotide (phosphate)
CDO	cumene dioxygenase	NaOH	sodium hydroxide
<i>cdw</i>	cell dry weight	NDO	naphthalene dioxygenase
COSY	correlation spectroscopy	NMO	<i>N</i> -methylmorpholine oxide
<i>cww</i>	cell wet weight	NMR	nuclear magnetic resonance
DAD	diode array detector	n.d.	not determined
DCM	dichloromethane	OD	optical density
<i>de</i>	diastereomeric excess	Os	osmium
DHAP	dihydroxyacetone phosphate	<i>P. putida</i>	<i>Pseudomonas putida</i>
DHQ	dihydroquinine	PCB	polychlorinated biphenyl
DHQD	dihydroquinidine	PCR	polymerase chain reaction
DMSO	dimethyl sulfoxide	PDB	protein data bank
DNA	deoxyribonucleic acid	PHAL	phthalazines
DO	dissolved oxygen	PLP	pyridoxal-5'-phosphate
ECN	effective carbon number	PMP	pyridoxamine-5'-phosphate
<i>E. coli</i>	<i>Escherichia coli</i>	PPD	1-phenylpropane-1,2-diol
<i>ee</i>	enantiomeric excess	ppm	parts per million
EH	epoxide hydrolase	PYR	pyrimidines
EI	electron ionization	RADH	ADH from <i>Ralstonia</i> sp.
epPCR	error-prone PCR	RF	relative response factor
Et <sub>2</sub> O	diethyl ether	RID	refractive index detector
EtOAc	ethyl acetate	RO	Rieske non-heme iron oxygenase
eV	electron volt	Ru	ruthenium
FAD	flavin adenine dinucleotide	s	second
FMN	flavin mononucleotide	SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
FruA	D-fructose 1,6-bisphosphate aldolase	SOC	super optimal broth with catabolite repression
g	gram	StEP	staggered extension process
GC	gas chromatography	TB	terrific broth
GC-FID	gas chromatography coupled to flame ionization detection	<i>t</i> -BuOH	<i>tert</i> -butanol
GC-MS	gas chromatography coupled to mass spectrometry	TCE	trichloroethylene
h	hour	TDO	toluene dioxygenase
HMBC	heteronuclear multiple bond correlation	ThDP	thiamine diphosphate
HPLC	high-performance liquid chromatography	TLC	thin-layer chromatography
HPP	2-hydroxy-1-phenylpropan-1-one	TMS	tetramethylsilane
HSQC	heteronuclear single quantum coherence	UV	ultra violet
HTS	high-throughput screening	<i>v/v</i>	volume/volume
IND	indoline	<i>vvm</i>	gas volume flow per unit of liquid volume per minute
IPTG	isopropyl β-D-1-thiogalactopyranoside	wt	wild type
		ω-TA	ω-transaminase
		<i>Y<sub>P/S</sub></i>	product yield coefficient

## ABSTRACT

The introduction of chirality into C=C double bonds is of special interest in organic synthesis. In particular, the catalytic asymmetric dihydroxylation (AD) of alkenes has attracted considerable attention due to the facile transformation of the chiral diol products into valuable derivatives. By chemical means, the metal-catalyzed AD of olefins provides both stereo- and regioselective *cis*-diol moieties. Next to their toxicity, however, these metal catalysts can also lead to byproduct formation as a result of oxidative fission.

In nature, Rieske non-heme iron oxygenases (ROs) represent promising biocatalysts for this reaction since they are the only enzymes known to catalyze the stereoselective formation of vicinal *cis*-diols in one step. ROs are key enzymes in the degradation of aromatic hydrocarbons and can target a wide variety of different arenes. Despite their broad substrate scope, limited data is available for the conversion of unnatural substrates by this class of enzymes. To explore their potential for alkene oxidation, three ROs were tested for the oxyfunctionalization of a set of structurally diverse olefins including linear and cyclic arene-substituted alkenes, cycloalkenes as well as several terpenes. Naphthalene- (NDO), benzene- (BDO) and cumene dioxygenases (CDO) from different *Pseudomonas* strains were selected as they are amongst the RO enzymes that have already been reported to catalyze the oxidation of a small number of olefins. The majority of compounds from the selected substrate panel could be converted by NDO, BDO or CDO and products were either isolated and identified by NMR analysis or using the authentic standards. Dependent on the substrate, allylic monohydroxylation was found in addition to the corresponding diol products, a reaction which is chemically still most reliably achieved by the use of SeO<sub>2</sub> in stoichiometric amounts.

However, having been evolved for the dihydroxylation of aromatic compounds, wild type ROs displayed low conversions (< 50%) and modest stereoselectivities ( $\leq 80\%$  *ee/de*) for several of the tested olefins. To overcome these limitations, changes in the active site topology of RO catalysts were introduced. A single targeted point mutation that was identified based on sequence and structural comparisons with other members of the RO family proved to be sufficient to generate BDO and CDO variants displaying remarkable changes in regio- and stereoselectivity for various substrates. In particular biotransformations with CDO M232A gave excellent stereoselectivities ( $\geq 95\%$  *ee/de*) and good activities (> 90%) also for linear alkenes, which have been reported to be challenging substrates for RO-catalyzed oxyfunctionalizations.

Site-saturation mutagenesis at position 232 in CDO revealed a correlation between the steric demand of the amino acid side chain and its influence on regio- and/ or stereoselectivities for styrene and indene. While the wild type enzyme almost exclusively catalyzed the dihydroxylation of the aromatic ring, the regioselectivity was shifted with decreasing side chain size to the terminal vinyl group of styrene, yielding up to 96% of the alkene-1,2-diol. For *cis*-1,2-indandiol formation, enantiocomplementary enzymes could be generated, a fact further highlighting the importance of position 232 for the engineering of ROs. Moreover, site-saturation mutagenesis of additional residues in the substrate binding pocket of CDO (F278, I288, I336 and F378) identified further positions having an influence on selectivity and product formation for alkene oxidation.

To proof the applicability of ROs for organic synthesis, semi-preparative scale biotransformations (70 mg) of selected substrates were performed with CDO M232A. Without further optimization of the reaction set-up, products were successfully isolated in > 30% yield. In addition, up-scaling of (*R*)-limonene hydroxylation to 4 L in a bioreactor with growing cells gave final isolated product titers of 0.4 g L<sup>-1</sup> even though substrate volatility and product toxicity diminished the yield.

In conclusion, these examples demonstrated that a single point mutation was sufficient to transform CDO wild type into an efficient catalyst, furthermore constituting the first example of the rational engineering of CDO and BDO enzymes for the oxyfunctionalization of a broad range of alkenes.

## ZUSAMMENFASSUNG

Die Einführung von Chiralität in C=C Doppelbindungen ist von besonderem Interesse in der organischen Synthese. Insbesondere die katalytische asymmetrische Dihydroxylierung (AD) von Alkenen hat aufgrund der einfachen Umwandlung der chiralen Diol-Produkte in wertvolle Verbindungen beträchtliche Aufmerksamkeit auf sich gezogen. In der Chemie liefert die Metall-katalysierte AD von Olefinen sowohl regio- als auch stereospezifisch *cis*-Diolverbindungen. Neben ihrer Toxizität können diese Metallkatalysatoren durch oxidative Spaltung jedoch auch zur Bildung von Nebenprodukten führen.

In der Natur stellen Rieske Nicht-Häm Eisen Oxygenasen (ROs) vielversprechende Biokatalysatoren für diese Reaktion dar, da sie die einzigen bekannten Enzyme sind, die die stereoselektive Bildung von *cis*-Diolen in nur einem Schritt ermöglichen. ROs sind Schlüsselenzyme beim Abbau von Aromaten und akzeptieren eine große Vielzahl von verschiedenen Verbindungen. Trotz ihres breiten Substratspektrums stehen nur begrenzte Informationen über die Oxidation von nicht natürlichen Substraten mit dieser Enzymklasse zur Verfügung. Um ihr Potential für die Umsetzung von Alkenen zu erforschen, wurden drei ROs für die Oxyfunktionalisierung eines Sets an strukturell unterschiedlichen Olefinen getestet, das sowohl lineare und zyklische Aryl-substituierte Alkene, Cycloalkene als auch verschiedene Terpene umfasste. Naphthalen- (NDO), Benzen- (BDO) und Cumoldioxygenasen (CDO) aus verschiedenen *Pseudomonas* Stämmen wurden ausgewählt, da sie zu den RO Enzymen gehören, mit denen bereits die Oxidation einiger weniger Olefine gezeigt werden konnte. Die Mehrheit der getesteten Verbindungen aus dem Substratpanel konnte von NDO, BDO oder CDO umgesetzt werden und die gebildeten Produkte wurden entweder isoliert und mittels NMR Spektroskopie charakterisiert oder anhand eines Standards identifiziert. Abhängig vom Substrat wurde neben den entsprechenden Diol-Produkten auch die allylische Monohydroxylierung beobachtet, eine Reaktion, die chemisch immer noch am verlässlichsten mit stöchiometrischen Mengen an SeO<sub>2</sub> durchgeführt wird.

Da die Wildtyp ROs von Natur aus die Dihydroxylierung aromatischer Verbindungen katalysieren, zeigten sie jedoch nur geringe Umsätze (< 50%) und moderaten Stereoselektivitäten ( $\leq 80\%$  *ee/de*) für mehrere der untersuchten Alkene. Um diese Einschränkungen zu überwinden, wurden Veränderungen in der Struktur des aktiven Zentrums der ROs eingeführt. Eine einzelne gerichtete Punktmutation, die anhand von Sequenz- und Strukturvergleichen mit anderen Mitgliedern der RO Familie identifiziert wurde, reichte aus, um CDO und BDO Varianten mit erheblichen Veränderungen in der Regio- und Stereoselektivität für verschiedene Substrate zu generieren. Insbesondere in Biotransformationen mit der CDO Mutante M232A konnten exzellente Stereoselektivitäten ( $\geq 95\%$  *ee/de*) und gute Aktivitäten (> 90%) auch für lineare Alkene erzielt werden, die als schwierige Substrate für die RO-katalysierte Oxyfunktionalisierung beschrieben worden sind.

Ortsgerichtete Sättigungsmutagenese an Position 232 in CDO zeigte einen Zusammenhang zwischen dem räumlichen Anspruch der Aminosäureseitenkette und ihrem Einfluss auf die Regio- und/ oder Stereoselektivität für zwei Substrate, Styrol und Inden. Während das Wildtyp Enzym fast ausschließlich die Dihydroxylierung des aromatischen Rings katalysierte, wurde die Regioselektivität mit abnehmender Größe der Seitenkette zur terminalen Vinylgruppe von Styrol verschoben und bis zu

96% des Alken-1,2-diols gebildet. Für die Bildung von *cis*-1,2-Indandiol konnten enantiokomplementäre Enzyme generiert werden, was die Bedeutung von Position 232 für das Engineering von ROs weiter verdeutlicht. Darüber hinaus wurden durch ortsgerichtete Sättigungsmutagenese von zusätzlichen Aminosäuren in der Substratbindetasche von CDO (F278, I288, I336 und F378) weitere Positionen identifiziert, die einen Einfluss auf die Selektivität und Produktbildung bei der Oxidation von Alkenen haben.

Um die Anwendbarkeit von ROs in der organischen Synthese zu zeigen, wurden Biotransformationen in einem semi-präparativen Maßstab (70 mg) mit ausgewählten Substraten und CDO M232A als Biokatalysator durchgeführt. Ohne weitere Optimierung des Reaktions-Setups konnten die entsprechenden Produkte mit > 30% Ausbeute erfolgreich isoliert werden. Zusätzlich wurde die Hydroxylierung von (*R*)-Limonen mit wachsenden Zellen in einem Bioreaktor auf 4 L hochskaliert, wobei die isolierte Produktausbeute von 0.4 g L<sup>-1</sup> durch die Flüchtigkeit des Substrats sowie die Produkttoxizität verringert wurde.

Schlussfolgernd zeigen diese Beispiele, dass eine einzelne Punktmutation ausreichend ist, um CDO Wildtyp in einen effizienten Biokatalysator zu verwandeln und stellen des Weiteren das erste Beispiel für das rationale Engineering von CDO und BDO Enzymen für die Oxyfunktionalisierung eines breiten Spektrums an Alkenen dar.

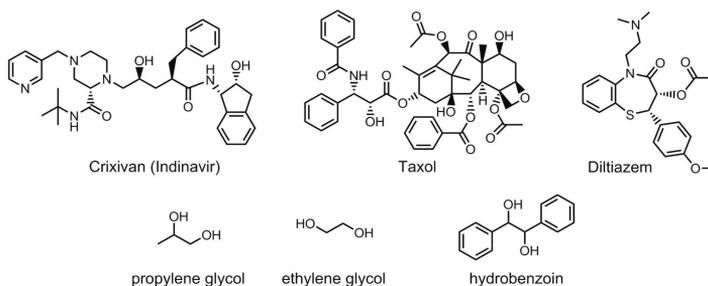
## 1. INTRODUCTION

### 1.1. Vicinal diols - Building blocks in the chemical and pharmaceutical industry

The introduction of chirality into C=C bonds is of special interest in organic synthesis.<sup>1</sup> In particular, the stereo- and regioselective oxidative functionalization of olefins which is one of the most challenging reactions in organic chemistry has attracted considerable attention.<sup>2-4</sup> Amongst the various products that can be generated, vicinal diols are important building blocks for the pharmaceutical and chemical industry.<sup>5</sup> Due to their facile transformation into various other functional groups, optically active diols represent one of the most sought after molecules in the area of asymmetric synthesis.<sup>6</sup>

Frequently, vicinal diols are used as synthons for the preparation of a variety of chiral compounds by converting one of the two hydroxyl groups into a leaving group that can be displaced by a nucleophile.<sup>7</sup> The preparation of 1,2-amino alcohols, diamines, diphosphines, sulfonates, halohydrin esters and epoxides has been reported.<sup>6,7</sup> Amongst the most important synthons for asymmetric synthesis are electrophilic building blocks including chiral epoxides and cyclic sulfates as well as halohydrins and glyceraldehyde derivatives. Their electrophilic nature facilitates bond-forming transformations and hence, these compounds are widely applied in the synthesis of natural products and biologically active molecules.<sup>7</sup> Examples for the preparation of pharmaceuticals containing 1,2-diol building blocks include the application of chiral cinnamate diols for the synthesis of diltiazem, an anti-angina drug, and Taxol side chains which was reported by Choudary and coworkers (Scheme 1).<sup>8</sup> In addition, a potential route to Crixivan® (Indinavir), a HIV protease inhibitor, involves *cis*-(1*S*,2*R*)-indandiol as synthon for the key intermediate *cis*-aminoindanol (Scheme 1).<sup>9</sup> Furthermore, enantiopure diols are frequently employed as chiral ligands and auxiliaries in asymmetric reactions.<sup>7</sup> Especially hydrobenzoin and its derivatives play important roles in organic synthesis (Scheme 1). Next to its application in asymmetric Michael reactions or cyclopropanations, hydrobenzoin can be further transformed into stilbenediamine which is employed for the preparation of chiral mediators for asymmetric Diels-Alder, aldol and allylmatalation reactions.<sup>7</sup>

Besides their major importance in organic synthesis, vicinal diols are also represented in the industrial production of bulk and fine chemicals.<sup>10</sup> Simple diols like propylene glycol and ethylene glycol are applied as polyester monomers or anti-freeze agents and produced on a multi ton scale per year while 1,2-hexanediol or 2,3-butanediol are important starting materials for the fine chemical industry.<sup>11</sup>



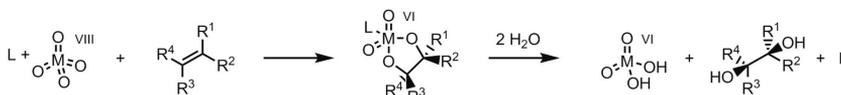
**Scheme 1:** Biologically active compounds containing diol moieties as building blocks (Crixivan, Taxol, Diltiazem), diols used as chiral ligands (hydrobenzoin) or industrially important diols like propylene and ethylene glycol.

## 1.2. Synthesis of vicinal *cis*-diols

With respect to their important role in organic synthesis, diverse methodologies have been developed to gain access to vicinal diols, focusing on selective catalysts that accept a broad range of substrates.<sup>3</sup> For the preparation of *trans*- and *cis*-1,2-diols, both chemical and enzymatic routes have been described in literature to obtain the oxidation product in an optically pure form. Herein, the synthesis of vicinal *trans*-diols employs selective reduction, aldol condensation or hydrolysis, either by chemical methods (reduction of aliphatic  $\alpha$ -hydroxy ketones with  $\text{Zn}(\text{BH}_4)_2$ , L-proline-catalyzed aldol condensation or hydrolysis of epoxides)<sup>12–14</sup> or biocatalytic means (epoxide hydrolases (EHs), aldolases and alcohol dehydrogenases (ADHs)),<sup>3,15,16</sup> In contrast, the chemical synthesis of *cis*-diols is generally heavy metal-catalyzed<sup>5</sup> and the direct addition of hydroxyl groups to olefins for the selective formation of vicinal *cis*-diols is an important reaction in asymmetric synthesis.<sup>17</sup> While metal oxide-based methodologies are well established,<sup>7</sup> also biocatalytic systems for the selective generation of vicinal *cis*-diols are available which mainly include oxidoreductases and lyases.<sup>18–21</sup>

### 1.2.1. Chemical catalysts: Metal-catalyzed asymmetric dihydroxylation (AD)

For direct *cis*-diol formation by chemical methods, heavy metal oxides are the most commonly employed catalysts in organic synthesis (Scheme 2).<sup>5,11</sup> The majority of published results within recent years are based on osmium tetroxide ( $\text{OsO}_4$ ),<sup>4</sup> however, also examples for the use of other metal oxides exist which include ruthenium, manganese or iron-oxo catalysts.<sup>5</sup> For metal-free systems (e.g. Prévost-Woodward-mediated dihydroxylation) or dihydroxylation based on molybdenum, palladium, cerium or technetium, the reader is referred to literature as the detailed reviewing of these methods would be beyond the scope of this thesis.<sup>5</sup>



**Scheme 2:** Metal-catalyzed dihydroxylation of alkenes. M = Os, Ru or Mn; L = ligand, adapted from Strassner.<sup>22</sup>

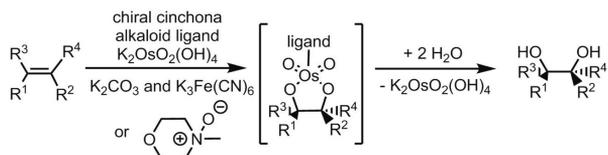
#### 1.2.1.1. Osmium-catalyzed AD (Sharpless AD)

One of the most widely applied techniques for the synthesis of vicinal *cis*-diols is the osmium(VIII)-catalyzed AD of alkenes developed by Sharpless and coworkers that has found numerous applications in organic chemistry.<sup>4</sup> The Sharpless AD was the first chemical reaction to form vicinal *cis*-diols with high enantioselectivity and still plays a key role in organic synthesis due to its high tolerance towards functional groups, good yields, and excellent stereospecificities.<sup>5</sup>

In the process, the dihydroxylation of an olefin substrate is catalyzed by  $\text{OsO}_4$  in the presence of a chiral cinchona alkaloid ligand (quinine or quinidine) and a stoichiometric secondary oxidant like potassium ferricyanide ( $\text{K}_3\text{Fe}(\text{CN})_6$ ) in combination with potassium carbonate ( $\text{K}_2\text{CO}_3$ ) or *N*-methylmorpholine oxide (NMO) for catalyst regeneration (Scheme 3).<sup>7</sup> During the reaction, coordination of the cinchona alkaloid to the osmium catalyst promotes the formation of a chiral complex with the amine ligand assumed to provide an enzyme-like binding pocket. This complex can

distinguish between the prochiral faces of an alkene substrate, leading to the stereoselective formation of vicinal *cis*-diols.<sup>23</sup>

Even though more than 400 different alkaloids have been tested for the AD of alkenes, the cinchona alkaloid derivatives developed by Sharpless *et al.* still represent the most effective approach under catalytic conditions.<sup>3</sup> Using these alkaloid derivatives, Sharpless and coworkers have designed a chiral ligand system consisting of the naturally derived dihydroquinine (DHQ) or dihydroquinidine (DHQD) units linked by a spacer (Scheme 4A).<sup>23</sup> In combination with DHQ and DHQD, the nature of these linkers can significantly influence the stereoselectivity of the AD.<sup>4</sup> Only three different classes of cinchona-based ligands proved to be sufficient for the dihydroxylation of a broad range of olefins. The most general ligands include phthalazines (PHAL) and pyrimidines (PYR) as linkers which afford high stereoselectivities for five of the six olefin classes (mono-, *trans*-di-, *gem*-di-, tri- and tetra-substituted olefins, Table 1).<sup>7</sup> For the dihydroxylation of *cis*-olefins, special indoline ligands (IND) are required. However, even when IND linkers are employed, stereoselectivities for *cis*-di-substituted alkenes are only low to moderate with *ee* values < 90%.<sup>24</sup>

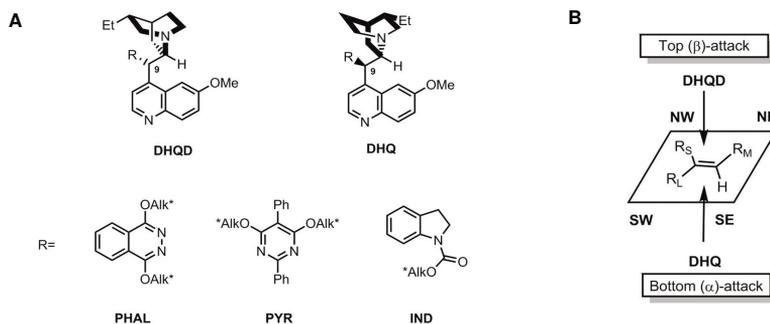


**Scheme 3:** The asymmetric dihydroxylation (AD) of olefins developed by Sharpless and coworkers. The reaction is catalyzed by osmium tetroxide (OsO<sub>4</sub>) in the presence of a chiral cinchona alkaloid ligand (quinine or quinidine) and a stoichiometric secondary reoxidant (K<sub>3</sub>Fe(CN)<sub>6</sub>/K<sub>2</sub>CO<sub>3</sub> or NMO).

**Table 1:** Preferred ligands for the AD of the different olefin classes I-VI with obtained stereoselectivities modified after Sharpless *et al.*<sup>7</sup>

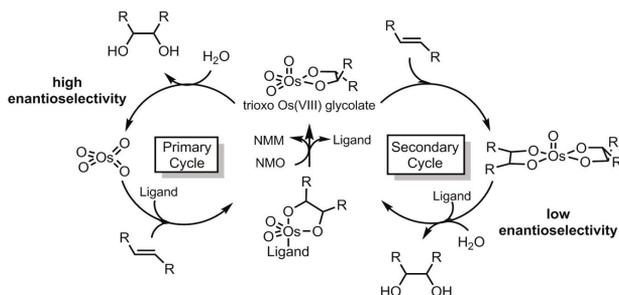
	I	II	III	IV	V	VI
Olefin class	mono	<i>gem</i> -di	<i>cis</i> -di	<i>trans</i> -di	tri	tetra
Best ligand	PYR/ PHAL	PHAL	IND	PHAL	PHAL	PYR/ PHAL
<i>ee</i> [%]	30-97	70-97	20-80	90-99.8	90-99	20-97

Using DHQ- and DHQD-derived ligands, both enantiomers of a diol are accessible during the hydroxylation reaction.<sup>3</sup> Even though being diastereomers,<sup>7</sup> the two cinchona alkaloids behave like enantiomers since mainly the nature of the substituent at C-9 of the alkaloid backbone is crucial for enantio-differentiation in *cis*-dihydroxylation (Scheme 4A).<sup>23</sup> The enantioselectivity in the AD reaction can be predicted dependent on the alkaloid ligand. With DHQD derivatives, the olefin will be attacked from the top face ( $\beta$ -face) whereas DHQ-based ligands promote attack from the bottom face ( $\alpha$ -face, Scheme 4B). However, this prediction does not apply for all alkenes and it can be ambiguous which group of the olefin is the large substituent (R<sub>L</sub>) and which the medium (R<sub>M</sub>) one to orient the substrate.<sup>25</sup>



**Scheme 4:** (A) Commonly used linkers PHAL, PYR and IND in combination with DHQD or DHQ (Alk<sup>\*</sup>). PHAL can be used for most alkene classes while PYR is best for mono- and tetra-substituted olefins. IND is applied for *cis*-di-substituted alkenes.<sup>7,25</sup> (B) Mnemonic device for the prediction of enantioselectivity. The olefin is oriented so that the large, medium and small substituents match best with the R<sub>L</sub>, R<sub>M</sub> and R<sub>S</sub> positions. R<sub>L</sub>: large, R<sub>M</sub>: medium, R<sub>S</sub>: small substituent. NW: north west, NE: north east, SW: south west, SE: south east quadrant.<sup>25</sup>

Since its development by Sharpless and coworkers, the osmium-catalyzed AD has been extensively optimized, not only in terms of ligand design but also regarding reaction conditions.<sup>3</sup> The reaction is based on the ligand acceleration effect<sup>7</sup> and experimental as well as theoretical work indicate that it presumably proceeds through a concerted [3+2]-like pathway which involves an osmium glycolate as intermediate.<sup>26</sup> The use of K<sub>3</sub>Fe(CN)<sub>6</sub> as oxidant in a two-phase system substantially improved enantioselectivities compared to NMO for osmium(VI) regeneration.<sup>27</sup> Under homogenous conditions as used for the process employing NMO, the reoxidant has constant access to all catalytic intermediates and the osmate(VI) monoglycolate undergoes reoxidation and can react with a second olefin, resulting in the formation of osmium(VI) bisglycolate. This complex can dihydroxylate the substrate albeit with low enantioselectivity as the chiral alkaloid ligand is not present (secondary cycle, Scheme 5).<sup>4,28</sup> With K<sub>3</sub>Fe(CN)<sub>6</sub> applied in biphasic reaction conditions, this secondary cycle was nearly completely avoided, increasing the enantioselectivity of the reaction.<sup>29</sup> Furthermore, under biphasic reaction conditions, the rate-limiting hydrolysis step of the osmium(VI) glycolate ester can be accelerated significantly by methanesulfonamide (MeSO<sub>2</sub>NH<sub>2</sub>). Thus, addition of MeSO<sub>2</sub>NH<sub>2</sub> allows oxidation of sterically hindered alkenes including tetra-substituted olefins as well as reaction temperatures of 0°C instead of room temperature, having also beneficial effects on selectivities.<sup>29,30</sup>



**Scheme 5:** Primary and secondary catalytic cycle under homogenous conditions using NMO as reoxidant.<sup>7</sup>

The commercial availability of  $\text{K}_2\text{OsO}_2(\text{OH})_4$  as non-volatile osmium source premixed with  $\text{K}_3\text{Fe}(\text{CN})_6/\text{K}_2\text{CO}_3$  and the chiral ligand as AD-mix  $\alpha$  ( $(\text{DHQ})_2\text{PHAL}$ ) or AD-mix  $\beta$  ( $(\text{DHQD})_2\text{PHAL}$ ) makes the reaction easily to be performed. Both AD-mixes can be employed for the dihydroxylation of mono-, *trans*-di-, *gem*-di- and tri-substituted olefins which are standard substrates for Sharpless AD and require very similar reaction conditions.<sup>7</sup> However, due to its electrophilic nature,  $\text{OsO}_4$  reacts only slowly with electron-deficient alkenes and hence, their stereoselective *cis*-dihydroxylation remains a major challenge.<sup>7,31</sup>

### 1.2.1.2. Osmium-free systems for the direct *cis*-dihydroxylation of olefins

Because of its efficiency, high functional group tolerance and excellent stereoselectivities, Sharpless AD is still the method of choice for the formation of enantiopure *cis*-diols and its importance was demonstrated with receiving the Nobel Prize 2001 in Chemistry.<sup>32</sup> However, regarding the high costs, volatility and toxicity of  $\text{OsO}_4$  that prevent a successful application on an industrial scale,<sup>33</sup> alternative oxidants were tested to circumvent the use of osmium.

Two alternative systems that have been applied for the *cis*-dihydroxylation of olefins include permanganate ( $\text{MnO}_4^-$ ) and ruthenium tetroxide ( $\text{RuO}_4$ ). Due to their lower toxicity, these oxidants provide a promising approach for the replacement of  $\text{OsO}_4$ .<sup>17,34</sup> Yet, efficient asymmetric ruthenium and manganese catalysts for dihydroxylation have not been reported to date as reactions are often not very selective and difficult to control. The high oxidative potential of both metal oxides impedes the termination of the oxidation reaction at the diol stage and overoxidation as well as formation of fission products are common side reactions.<sup>17,33</sup> Thus, in particular  $\text{RuO}_4$  is more commonly employed for oxidative cleavage and in ketohydroxylations than for the formation of vicinal *cis*-diols.<sup>5,17,33</sup>

Similar to Sharpless AD, engineering of the process conditions led to improvements in the dihydroxylation efficiency for both oxidants. For  $\text{RuO}_4$ , high catalyst loadings (7 mol%) which offset the lower price of ruthenium compared to osmium could be avoided under acidic conditions using  $\text{H}_2\text{SO}_4$  or  $\text{CeCl}_3$  to accelerate the reaction.<sup>35,36</sup> In the latter system, glycol cleavage is reduced compared to the Brønsted acid-accelerated method and hence, the reaction time can be increased, allowing the conversion of a wider scope of olefins in good to excellent yields.<sup>36</sup> Yet, despite these improvements, the  $\text{RuO}_4$ -catalyzed dihydroxylation is still not as general applicable for a broad range of alkenes as Sharpless AD.<sup>17</sup> Using  $\text{MnO}_4^-$  as oxidant which reacts with alcohols, alkenes, aldehydes, saturated C-H bonds as well as other functional groups,<sup>34</sup> the reaction conditions have to be controlled carefully to avoid extensive side-product formation and the reaction pathway is influenced by solvent, pH and substrate. The best conditions for *cis*-dihydroxylation are provided in alkaline media while formation of  $\alpha$ -hydroxy ketones is catalyzed under neutral or slightly basic conditions. However, reactions are always accompanied by C-C bond cleavage which predominates under acidic conditions.<sup>34</sup> As the use of  $\text{MnO}_4^-$  in stoichiometric amounts leads to significant manganese dioxide ( $\text{MnO}_2$ ) by-product formation,<sup>37</sup> different methods for catalytic manganese dihydroxylations have been developed to avoid problems caused by  $\text{MnO}_2$  removal.<sup>38,39</sup>

To date, only a limited number of AD reactions catalyzed by  $\text{RuO}_4$  or  $\text{MnO}_4^-$  have been reported in literature. For  $\text{RuO}_4$ , these were all based on the chiral auxiliary approach with moderate to good overall yields and good to excellent enantio- or diastereoselectivities.<sup>40,41</sup> Furthermore, the two

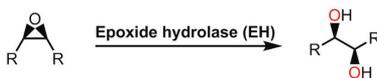
methods currently described for the AD using manganese resulted in moderate to good stereoselectivities.<sup>42,43</sup>

Besides RuO<sub>4</sub> and MnO<sub>4</sub><sup>-</sup>, the increasing interest in the use of iron complexes for catalysis due to their abundance in nature as well as their lack of toxicity has prompted the development and design of non-heme iron catalysts for olefin *cis*-dihydroxylation.<sup>11,44</sup> First examples of bio-inspired non-heme iron catalysts have been reported by Que and coworkers using tetradentate nitrogen-donor ligands with topologies that allow two *cis*-oriented coordination sites to be available for peroxide binding and activation which is used as oxidant. These iron complexes catalyzed olefin *cis*-dihydroxylation and epoxidation although yields and turnovers were generally low.<sup>45</sup> To achieve higher ratios of diol:epoxide, the group of Que successfully modulated the complexes to obtain a ligand environment that more closely mimics the environment of the mononuclear iron center in Rieske non-heme iron oxygenase (RO) enzymes.<sup>46</sup> Currently, there are only two examples known in literature for the AD with iron complexes using hydrogen peroxide (H<sub>2</sub>O<sub>2</sub>) as reoxidant which gave moderate to excellent enantioselectivities (> 97% *ee*).<sup>47,48</sup> Yet, despite these promising results for the iron-catalyzed *cis*-dihydroxylation of olefins within the last years, the synthetic utility of bio-mimetic iron complexes has not been proven so far. Their application in organic synthesis has been hampered by the iron-mediated decomposition of H<sub>2</sub>O<sub>2</sub> which obstructs its efficient use as oxidant. The generation of highly reactive free hydroxyl radicals induces the degradation of ligand, substrate and product, usually requiring a high ratio of substrate to H<sub>2</sub>O<sub>2</sub>.<sup>11</sup>

### 1.2.2. Biocatalytic access to *cis*-1,2-diols

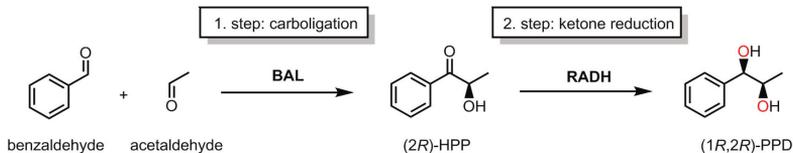
The development of sustainable, efficient and selective catalysts is one of the major goals in organic synthesis.<sup>49,50</sup> Metal-based reagents for the asymmetric *cis*-dihydroxylation of C=C double bonds are well established but the low sustainability of heavy metal oxides with regard to atom economy, toxicity, side-product formation and catalyst stoichiometry has prompted the search for alternative strategies.<sup>51</sup> In this context, biocatalytic routes for the selective formation of vicinal *cis*-diols have generated great interest due to the intrinsic high regio-, stereo- and chemoselectivity of many enzymatic reactions that are catalyzed under mild conditions.<sup>52</sup>

As an alternative to chemical means, several enzymatic routes to *cis*-1,2-diols have been explored. Herein, the use of EHs represent an exception as these enzymes usually catalyze epoxide hydrolysis to yield vicinal diols in *anti*-configuration.<sup>53</sup> Formation of *cis*-diols is rare and restricted to *meso*-epoxides. The first example has been described by Bellucci *et al.*, reporting the biocatalytic synthesis of (*R,R*)-(+)-1,2-diphenyl-1,2-ethanediol by microsomal EH from rabbit liver which catalyzes the ring opening of *cis*-stilbene oxide with high stereoselectivity.<sup>54</sup> In addition, Zhao and coworkers identified several microbial EHs which provide *cis*-1,2-diols from different aryl *meso*-epoxide substrates with good to excellent stereoselectivities (up to 99% *ee*) for the (*R,R*)-enantiomer (Scheme 6).<sup>55</sup>



**Scheme 6:** Epoxide hydrolase (EH)-catalyzed formation of vicinal *cis*-diols by hydrolysis of aryl *meso*-epoxides. R = C<sub>6</sub>H<sub>5</sub>, 2-Cl-C<sub>6</sub>H<sub>5</sub>, 2-F-C<sub>6</sub>H<sub>5</sub>, 3-Cl-C<sub>6</sub>H<sub>5</sub>, 4-Cl-C<sub>6</sub>H<sub>5</sub>, 2-pyridyl, 3-pyridyl or 4-pyridyl.<sup>55</sup>

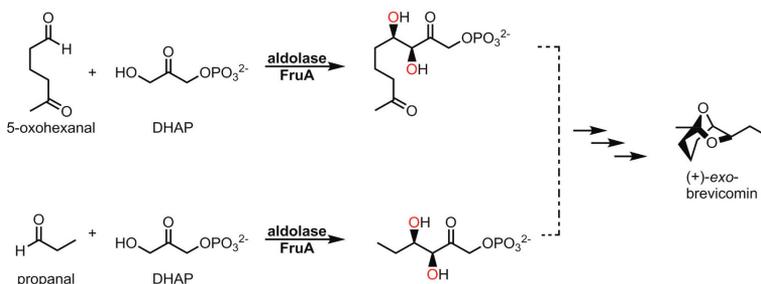
Apart from EHS, the main routes to vicinal *cis*-diols include the reduction of  $\alpha$ -diketones and  $\alpha$ -hydroxy ketones by ADHs, aldolase-catalyzed reactions and *cis*-dihydroxylation of C=C double bonds by ROs.<sup>18–21</sup> Amongst these biocatalysts, ADHs have been extensively studied during the last decade by both academia and industry. Due to their high stereoselectivities and availability, these enzymes represent the method of choice for the synthesis of chiral alcohols.<sup>56</sup> Yet, ADHs can also be applied for the formation of *cis*-1,2-diols starting from  $\alpha$ -diketones or  $\alpha$ -hydroxy ketones and examples include ADHs from *Rhodococcus ruber* DSM44541, *Lactobacillus brevis* and *Ralstonia* sp. as well as diacetyl reductase from *Bacillus stearothermophilus* and glycerol dehydrogenase from *Enterobacter aerogenes* or *Cellulomonas* sp..<sup>16,57–60</sup> Due to its excellent stereoselectivity for the reduction of various  $\alpha$ -hydroxy ketones, Jakoblinnert and Rother applied ADH from *Ralstonia* sp. (RADH) in the two-step synthesis of 1-phenylpropane-1,2-diol (Scheme 7).<sup>18</sup> Starting from cheap aldehydes, diol formation was accomplished with excellent selectivities (> 99% *ee/de*) and high product concentrations (space-time yields up to 327 g L<sup>-1</sup> d<sup>-1</sup>). In a first step, the thiamine diphosphate (ThDP)-dependent benzaldehyde lyase (BAL) from *Pseudomonas fluorescens* catalyzed the carbonylation of benzaldehyde and acetaldehyde yielding (*2R*)-2-hydroxy-1-phenylpropan-1-one ((*2R*)-HPP). (*2R*)-HPP was subsequently reduced by RADH giving (1*R*,2*R*)-1-phenylpropane-1,2-diol ((1*R*,2*R*)-PPD) as product. Employing this two-step approach, also the opposite enantiomer of PPD would be accessible by combining different ThDP-dependent enzymes and ADHs possessing Prelog or anti-Prelog specificities.<sup>18,52</sup>



**Scheme 7:** Enzymatic two-step cascade for the synthesis of the vicinal *cis*-diol (1*R*,2*R*)-1-phenylpropane-1,2-diol ((1*R*,2*R*)-PPD).<sup>18</sup> BAL: benzaldehyde lyase, RADH: ADH from *Ralstonia* sp..

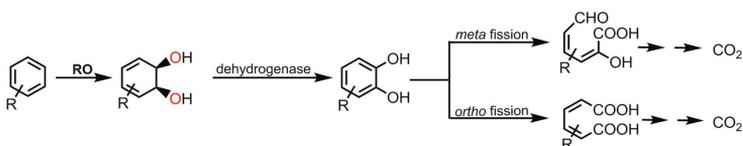
Besides ADHs, aldolases can be employed in the formation of 1,2-diol units. Dependent on the type of enzyme, *cis*-diols are generated during the aldolase-catalyzed carbonylation of two molecules. Dihydroxyacetone phosphate (DHAP)-dependent aldolases catalyze the C-C bond formation at 1,2-diol junctions that link a DHAP unit with a large variety of aldehyde acceptors. During the reaction,  $\alpha,\beta$ -disubstituted carbonyl derivatives are generated, containing two novel vicinal chiral centers at the new C-C bond.<sup>19</sup> Their stereochemistry can be directed by using different stereocomplementary DHAP-dependent aldolases for the reaction.<sup>15</sup> DHAP-dependent aldolases typically yield carbohydrates or carbohydrate-derived materials according to their substrates but can also be applied in the generation of *cis*-diol synthons for non-carbohydrate natural products. The DHAP-dependent aldolase D-fructose 1,6-bisphosphate aldolase (FruA) can catalyze the carbonylation of DHAP and 5-oxohexanal or propanal, yielding enantiopure vicinal *cis*-diol structures which represent potential intermediates in the synthesis of (+)-*exo*-brevicomine, a beetle pheromone (Scheme 8).<sup>19</sup> However, even though aldolases can generally tolerate a broad scope of aldehydes as acceptors, the donor

compound is often invariable and hence, only a limited structural diversity is accessible for *cis*-1,2-diols.<sup>61</sup> Furthermore, the reversibility of the catalyzed aldol reaction is an additional drawback.<sup>16</sup>



**Scheme 8:** D-Fructose 1,6-bisphosphate aldolase (FruA)-catalyzed generation of different precursors for the chemoenzymatic synthesis of (+)-*exo*-brevicimin, adapted from Fessner.<sup>19</sup>

One of the most promising biocatalytic approaches for the selective formation of vicinal *cis*-diols is provided by ROs. In contrast to EHS, ADHs and aldolases which catalyze the formation of both *trans*- and *cis*-diols,<sup>53–55,58,61</sup> these oxygenases stereoselectively introduce two hydroxyl groups in one enzymatic step solely in a *cis*-directed fashion.<sup>62,63</sup> Among the vast number of hydroxylating enzymes, as far as known, only ROs possess this unique ability, making them highly interesting biocatalysts for organic synthesis.<sup>64</sup> By the incorporation of both atoms from O<sub>2</sub> into the aromatic ring, ROs, also called aromatic-ring-hydroxylating dioxygenases, initiate bacterial pathways for the degradation of a wide range of xenobiotic compounds.<sup>65</sup> Their presence in arene-degrading soil bacteria was first established in the late 1960s by Gibson and coworkers when ROs were found in strains capable of growing on toluene, naphthalene or benzene as sole source of carbon.<sup>66,67</sup> The resonance energy stabilizing the aromatic ring structure can be overcome by the RO-catalyzed oxyfunctionalization, leading to the formation of a *cis*-dihydrodiol intermediate as a common first step which is further degraded by ring fission and subsequent breakdown into the intermediates of the tricarboxylic acid cycle (Scheme 9).<sup>68–71</sup> Like other enzymes involved in metabolic or detoxification pathways, ROs have a broad substrate scope and show a high stereospecificity for aromatic compounds.<sup>72</sup>

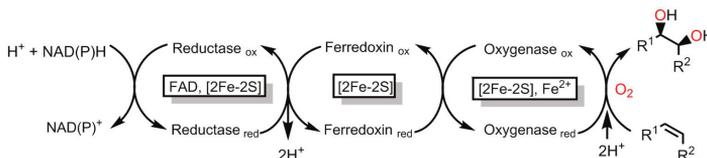


**Scheme 9:** Microbial degradation of aromatic compounds. In the first step, dihydroxylation of the aromatic nucleus is catalyzed by Rieske non-heme iron oxygenases (ROs).

Besides arene dihydroxylation, also a few examples are known for the RO-catalyzed oxyfunctionalization of aliphatic compounds.<sup>9,73–75</sup> Due to their substrate promiscuity, ROs were selected as biocatalysts for this project as they might provide a potential alternative to the heavy metal-catalyzed AD of alkenes.

### 1.3. Rieske non-heme iron oxygenases (ROs)

ROs are soluble multicomponent systems that harness the reductive power of NAD(P)H for oxygen activation. The cofactor-derived electrons are transferred *via* a reductase and, dependent on the enzyme, a ferredoxin to the terminal oxygenase component containing the active site (Scheme 10).<sup>76</sup> As ROs belong to the class of non-heme iron dependent enzymes, O<sub>2</sub> activation takes place at a mononuclear iron(II) center.<sup>77</sup> The divalent iron in the active site is coordinated by the so-called 2-His-1-carboxylate facial triad, a common structural motif consisting of two conserved histidines and a carboxylic acid (aspartate or glutamate). The motif provides a platform to anchor the iron while maintaining three *cis*-oriented sites available on the iron octahedron to coordinate other endogenous or exogenous ligands such as O<sub>2</sub>, substrate or cofactors.<sup>51,78</sup> Besides ROs, several members of the non-heme iron dependent enzyme superfamily employ the 2-His-1-carboxylate facial triad for iron coordination including  $\alpha$ -ketoglutarate-dependent dioxygenases, extradiol cleaving catechol dioxygenases, pterin-dependent hydroxylases and other oxidases like isopenicillin *N* synthase.<sup>79</sup> Although their active sites are based on unrelated protein folds and catalyze a wide range of different reactions, these enzymes share the ability to couple O<sub>2</sub> reduction with substrate oxyfunctionalization.<sup>80</sup> In addition to the non-heme iron center, the RO oxygenase component contains a second cofactor, the so-called Rieske [2Fe-2S] iron sulfur cluster that is involved in the electron transfer to the catalytic non-heme iron. In contrast to plant-type [2Fe-2S] centers being coordinated by four cysteine residues, the Rieske cluster is complexed by two conserved histidines and two conserved cysteines, resulting in significantly different spectroscopic properties and a more positive midpoint reduction potential.<sup>76,81</sup>



**Scheme 10:** Example of an RO electron transfer chain from NAD(P)H *via* a reductase and (dependent on the RO system) a ferredoxin to the terminal oxygenase component, modified after Parales and Resnick.<sup>82</sup>

To date, more than 100 ROs have been identified and completely sequenced.<sup>82</sup> Due to their versatility, ROs are considered as the non-heme analogue of cytochrome P450 monooxygenases and next to their relaxed substrate specificity, these enzymes can catalyze various oxidation reactions.<sup>20,82–84</sup>

#### 1.3.1. Classification of RO systems

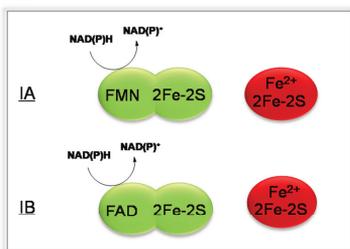
For the classification of ROs, different systems have been proposed which are either based on the nature of the electron transport components, the oxygenase sequence homology or a combination of both.<sup>64,71,85–88</sup> The historical system proposed by Batie *et al.* in 1991 classified ROs according to their electron transfer components and the nature of their redox centers into three main classes (Table 2).<sup>85</sup> Class I enzymes comprise two components, a flavin and [2Fe-2S] cluster containing reductase as well as an oxygenase. Class II and class III dioxygenases are both three component systems employing a ferredoxin for electron transfer between the reductase and oxygenase component. In class II, the

flavin and [2Fe-2S] redox centers are on separate proteins, whereas the reductase of class III contains both a flavin and a [2Fe-2S] cluster but also requires an additional [2Fe-2S] center on the ferredoxin for electron transfer to the terminal oxygenase. Further subdivision of class I and II into group A and B is based on the type of flavin (FMN or FAD, class I) in the reductase and the coordination of the [2Fe-2S] center in the ferredoxin (class II, Figure 1). In class IA, the reductase contains an N-terminal FMN whereas in class IB, a FAD is present as cofactor. Class IIA has a plant-type ferredoxin component while in class IIB, a Rieske-type iron-sulfur cluster is present in the ferredoxin (Figure 1).<sup>76</sup>

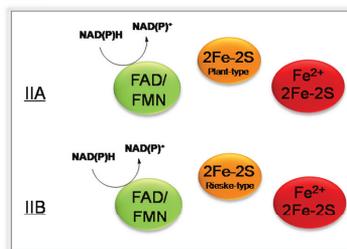
**Table 2:** Classification of different RO systems according to their components, adapted from Maison *et al.*<sup>76</sup>

System	Class	Reductase	Ferredoxin	Oxygenase	Example
2-component	IA	FMN, [2Fe-2S]	-	[2Fe-2S], Fe	Phthalate dioxygenase
	IB	FAD, [2Fe-2S]	-	[2Fe-2S], Fe	Benzoate dioxygenase
3-component	IIA	FAD	[2Fe-2S]	[2Fe-2S], Fe	Pyrazon dioxygenase
	IIB	FAD	[2Fe-2S]	[2Fe-2S], Fe	Benzene dioxygenase
					Toluene dioxygenase
III	FAD, [2Fe-2S]	[2Fe-2S]	[2Fe-2S], Fe	Naphthalene dioxygenase	

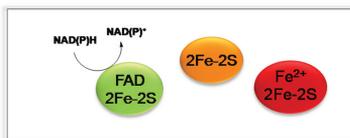
#### CLASS I



#### CLASS II



#### CLASS III



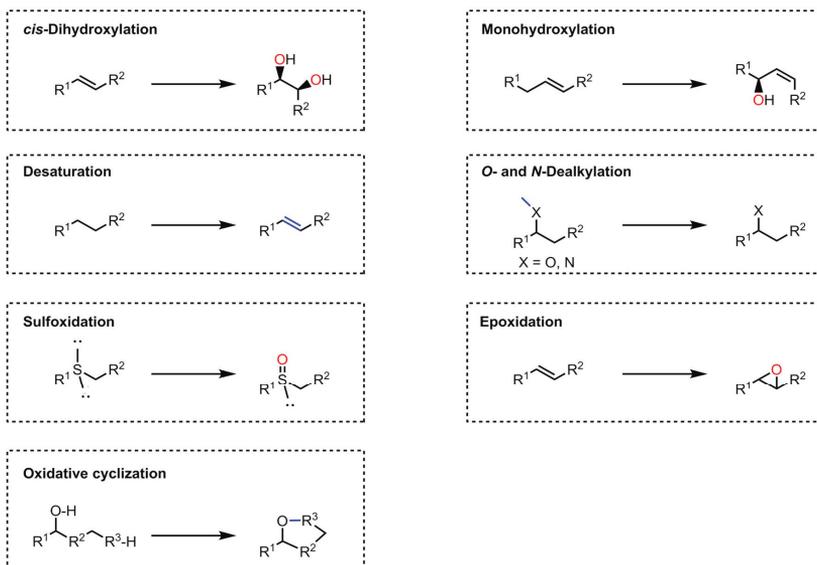
**Figure 1:** Classification of ROs by their components (green: oxygenase, red: reductase, orange: ferredoxin). Class I: two-component dioxygenases, flavin and iron-sulfur cluster are combined in the same protein, Class IA: FMN, Class IB: FAD. Class II: three-component dioxygenases, electron-transport chain with a flavoprotein and a separate ferredoxin, Class IIA: plant-type ferredoxin, Class IIB: Rieske-type iron-sulfur cluster in the ferredoxin. Class III: three-component dioxygenases, electron-transfer chain with iron-sulfur flavoprotein and ferredoxin.

Since its publication, the Batie classification scheme has been broadly applied. However, as this system is solely based on the electron transfer components, it is lacking any information about the terminal oxygenase catalyzing the dihydroxylation reaction.<sup>64</sup> Due to difficulties to integrate many newly identified oxygenases into the three classes, several other classification systems have been suggested that also include the oxygenase components.<sup>64,71,86–88</sup>

In 1996, the group of Werlen proposed a classification scheme based on amino acid sequence comparisons of oxygenase  $\alpha$ -subunits, identifying four dioxygenase families (naphthalene, toluene/benzene, biphenyl and benzoate/toluate families).<sup>86</sup> The clustering of dioxygenases into families correlates in general with the native substrates oxidized by the corresponding members. In addition to the former mentioned families, the phthalate family was identified by Gibson and Parales which is a large and diverse group of enzymes including both mono- and dioxygenases that oxidize aromatic acids.<sup>64</sup> However, there are also several oxygenases that do not cluster with any of these families and thus, in updated classification systems as proposed by Kweon or Chakraborty *et al.*, oxygenase sequence information as well as its functional properties were included.<sup>71,87</sup>

### 1.3.2. Substrate scope and reaction spectrum

ROs catalyze the asymmetric *cis*-dihydroxylation of C=C double bonds with high stereo- and regioselectivity. Their broad substrate spectrum comprises more than 300 substrates including monocyclic aromatic compounds, polycyclic and heterocyclic arenes, substituted aromatics, halogenated arenes as well as aromatic acids and has been investigated in detail for several enzymes.<sup>89</sup> Toluene dioxygenases (TDOs) from *Pseudomonas putida* (*P. putida*) F1 and *P. putida* UV4<sup>a</sup> catalyze the oxidation of more than 100 compounds and also NDO from *Pseudomonas* sp. NCIB 9816-4 has been reported to accept over 75 different substrates.<sup>82</sup> Furthermore, these versatile enzymes can catalyze various oxidation reactions including monohydroxylations, sulfoxidations, dealkylations, desaturations, oxidative cyclizations and epoxidations (Scheme 11).<sup>20,82–84</sup>



**Scheme 11:** Diverse reactions catalyzed by ROs including *cis*-dihydroxylation, allylic/ benzylic monohydroxylation, desaturation, *O*- and *N*-dealkylation, sulfoxidation, epoxidation and oxidative cyclization.

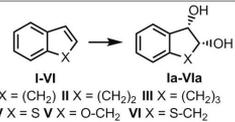
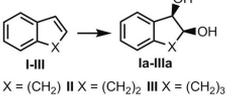
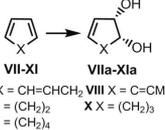
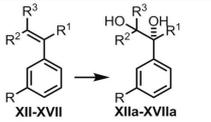
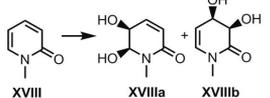
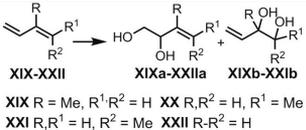
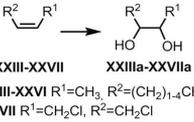
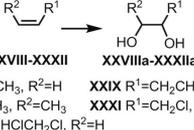
<sup>a</sup> *P. putida* UV4 is a mutant of *P. putida* NCIMB 11767 lacking *cis*-dihydrodiol dehydrogenase activity.<sup>222</sup>

Although there has been much work on the dihydroxylation of aromatic compounds, little has been reported on the conversion of unnatural substrates by ROs. To date, NDO, TDO, benzene (BDO) and cumene dioxygenase (CDO) enzymes from different *Pseudomonas* strains have been shown to catalyze the dihydroxylation of a small number of olefins (Table 3). Most of these studies were performed with benzocycloalkene or analogous substrates.<sup>20</sup> One of the first reported alkene dihydroxylation reactions was the TDO-catalyzed oxidation of indene **I** yielding the corresponding *cis*-diol (1*S*,2*R*)-**Ia** (Table 3) next to the allylic alcohol.<sup>90,91</sup> NDO from *Pseudomonas* sp. NCIB 9816-4 catalyzes the same reaction but favors the opposite enantiomer (1*R*,2*S*)-**Ia**.<sup>92</sup> With TDO and NDO, an opposite stereochemistry was also found for the corresponding *cis*-diols of 1,2-dihydronaphthalene **II** and benzocyclohept-1-ene **III** ((1*S*,2*R*)- and (1*R*,2*S*)-**Ila** as well as (3*S*,4*R*)- and (3*R*,4*S*)-**Illa**, Table 3),<sup>93</sup> making both enzymes enantiocomplementary biocatalysts.<sup>94</sup> TDO has been shown in addition to be able to dihydroxylate several heterocyclic alkenes (including substrates **IV-VI**) with high enantioselectivity to the corresponding *cis*-diols **Iva-VIa**.<sup>95</sup> Oxidation of heteromonocyclic alkenes was also reported for NDO and the dihydroxylation of the conjugated diene *N*-methyl-2-pyridone **XVIII** yielded the *cis*-diol products **XVIIIa** and **XVIIIb**.<sup>96</sup>

The dioxygenase-catalyzed formation of vicinal *cis*-diols from different conjugated monoalkenes and polyenes could be shown by Boyd and coworkers. Thereby, the asymmetric dihydroxylation of mono-substituted, *gem*-di-substituted, *cis*-di-substituted and tri-substituted olefins but not of *trans*-di- and tetra-substituted alkene bonds was reported.<sup>75</sup> In general, the preference for dihydroxylation of the conjugated alkene or arene group was dependent on the dioxygenase and alkene type. Several substituted styrene substrates **XII-XVII** were converted by TDO from *P. putida* UV4 in general to the corresponding arene-*cis*-1,2-dihydrodiols while the alkene-1,2-diols **XIIa-XVIIa** were formed to a lesser extent. In contrast, NDO from *P. putida* NCIMB 8859 which is unable to promote the dihydroxylation of most mono-substituted benzenes, could catalyze diol formation of **XII-XVII** without arene *cis*-dihydroxylation, yielding the alkene-1,2-diols.<sup>75</sup> Additionally, several non-aromatic conjugated cyclic dienes (C<sub>6</sub>-C<sub>8</sub>, **IX-XI**) and trienes (C<sub>7</sub>-C<sub>8</sub>, **VII-VIII**) of different ring sizes have been reported to undergo stereoselective *cis*-dihydroxylation catalyzed by TDO and NDO (Table 3).<sup>75,97</sup>

Beside the dihydroxylation of cyclic alkenes, the formation of *cis*-diols from acyclic olefins has been reported. However, acyclic alkenes are generally poor substrates for the RO-catalyzed dihydroxylation<sup>75</sup> and therefore, only few examples known. In 2000, Boyd and coworkers reported the biotransformation of isoprene **XIX** and related dienes (C<sub>4</sub> and C<sub>5</sub>, **XX-XXII**) to the vicinal *cis*-diols **XIXa,b-XXIIa,b** by *P. putida* ML2, a wild-type source of BDO.<sup>74</sup> Based on obtained product yields, *cis*-dihydroxylation seemed to occur in the sequence mono-substituted alkene > *cis*-di-substituted alkene > *gem*-di-substituted alkene > *trans*-di-substituted alkene. The yields of diols were relatively low (< 10%) and whereas the stereoselectivity is excellent for most cyclic substrates, it dropped for the open chain derivatives to max. 74% *ee*.<sup>74</sup> Furthermore, CDO and TDO have been shown to convert different halogenated and nonhalogenated acyclic alkenes **XXIII-XXXII** into the corresponding *cis*-diols.<sup>73,98</sup> For both dioxygenases, only olefins lacking chlorine substituents on double bond carbon atoms were dihydroxylated, while in the presence of chlorine atoms, monohydroxylation occurred.<sup>73,98</sup>

**Table 3:** Dioxygenase-catalyzed *cis*-dihydroxylation of cyclic and acyclic alkenes.<sup>20,74,75,91,93,95,97</sup> Monohydroxylated products or products resulting from the dihydroxylation of arene moieties are not shown.

Enzyme (strain)	Substrate/ Product	ee [%]	Yield [%]
TDO <sup>[a]</sup> ( <i>P. putida</i> UV4)	 I X = (CH <sub>2</sub> ) II X = (CH <sub>2</sub> ) <sub>2</sub> III X = (CH <sub>2</sub> ) <sub>3</sub> IV X = S V X = O-CH <sub>2</sub> VI X = S-CH <sub>2</sub>	Ia: 20 IIa-VIa: > 98	10-47
NDO <sup>[a]</sup> ( <i>Pseudomonas</i> sp. NCIB 9816-4)	 I X = (CH <sub>2</sub> ) II X = (CH <sub>2</sub> ) <sub>2</sub> III X = (CH <sub>2</sub> ) <sub>3</sub>	Ia: 72-80 <sup>[c]</sup> IIa-IIIa: > 98	50-65
NDO ( <i>P. putida</i> NCIMB 8859) TDO ( <i>P. putida</i> UV4)	 VII X = CH=CHCH <sub>2</sub> VIII X = C=CMe <sub>2</sub> IX X = (CH <sub>2</sub> ) <sub>2</sub> X X = (CH <sub>2</sub> ) <sub>3</sub> XI X = (CH <sub>2</sub> ) <sub>4</sub>	VIIa-XIa: ≥ 98	4-24 <sup>NDO</sup> / 4-29 <sup>TDO</sup>
NDO <sup>[a]</sup> ( <i>P. putida</i> NCIMB 8859) TDO <sup>[b]</sup> ( <i>P. putida</i> UV4)	 XII R-R <sup>3</sup> = H XIII R = Cl, R <sup>1</sup> -R <sup>3</sup> = H XIV R = Me, R <sup>1</sup> -R <sup>3</sup> = H XV R <sup>1</sup> = Me, R, R <sup>2</sup> , R <sup>3</sup> = H XVI R <sup>2</sup> = Me, R, R <sup>1</sup> , R <sup>3</sup> = H XVII R = F, R <sup>1</sup> -R <sup>3</sup> = H	XII: 80 <sup>NDO</sup> /88 <sup>TDO</sup> XIII: 56 <sup>NDO</sup> /42 <sup>TDO</sup> XIV: 66 <sup>NDO</sup> /48 <sup>TDO</sup> XV: 46 <sup>NDO</sup> / <sub>TDO</sub> XVI: 82 <sup>NDO</sup> / <sub>TDO XVII: 62<sup>NDO</sup>/62<sup>TDO</sup></sub>	12-60 <sup>NDO</sup> / 3-22 <sup>TDO</sup>
NDO ( <i>Pseudomonas</i> sp. NCIB 9816-4)	 XVIII XVIIIa XVIIIb	n.d.	55
BDO ( <i>P. putida</i> ML2)	 XIX R = Me, R <sup>1</sup> R <sup>2</sup> = H XX R,R <sup>2</sup> = H, R <sup>1</sup> = Me XXI R,R <sup>1</sup> = H, R <sup>2</sup> = Me XXII R-R <sup>2</sup> = H	XIXa/XIXb: 16/44 XXa/XXb: 38/33 XXIa/XXIb: 74/70 XXIIa: 25	< 10
TDO <sup>[a]</sup> ( <i>P. putida</i> UV4)	 XXIII-XXVII XXIIIa-XXVIIa XXIII-XXVI R <sup>1</sup> = CH <sub>3</sub> , R <sup>2</sup> = (CH <sub>2</sub> ) <sub>1-4</sub> CH <sub>3</sub> XXVII R <sup>1</sup> = CH <sub>2</sub> Cl, R <sup>2</sup> = CH <sub>2</sub> Cl	n.d.	n.d.
CDO <sup>[a]</sup> ( <i>P. fluorescens</i> IP01)	 XXVIII-XXXII XXVIIIa-XXXIIa XXVIII R <sup>1</sup> = CH <sub>3</sub> , R <sup>2</sup> = H XXIX R <sup>1</sup> = CH <sub>2</sub> CH <sub>3</sub> , R <sup>2</sup> = H XXX R <sup>1</sup> = CH <sub>3</sub> , R <sup>2</sup> = CH <sub>3</sub> XXXI R <sup>1</sup> = CH <sub>2</sub> Cl, R <sup>2</sup> = CH <sub>2</sub> Cl XXXII R <sup>1</sup> = CHClCH <sub>2</sub> Cl, R <sup>2</sup> = H	n.d.	n.d.

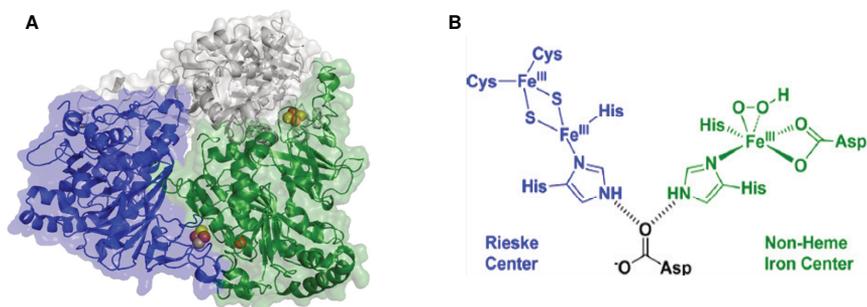
<sup>[a]</sup> Next to *cis*-dihydroxylation of the C=C double bond, also monohydroxylation in allylic position was observed.

<sup>[b]</sup> TDO also catalyzed the dihydroxylation of the aromatic ring forming arene-1,2-dihydrodiols. <sup>[c]</sup> ee Values vary dependent on the bacterial strain employed. n.d.: not determined

### 1.3.3. Structure and catalytic mechanism of ROs

In the RO multicomponent system, the terminal oxygenase subunit catalyzes substrate oxyfunctionalization. Dependent on the RO enzyme, the oxygenase either comprises two separate proteins, a large catalytic subunit ( $\alpha$ ) and a small subunit ( $\beta$ ) in hetero-multimeric form ( $\alpha_n\beta_n$ ), or lacks the  $\beta$ -subunit and exists in homo-multimeric form ( $\alpha_n$ ).<sup>63,71</sup> The  $\alpha$ -subunit contains the largely hydrophobic active site whereas the  $\beta$ -subunit is assumed to have a structural function. However, in distinct RO systems, the  $\beta$ -subunit was proposed to have an influence on substrate specificity.<sup>99</sup> For ROs being devoid of a  $\beta$ -subunit, loops on the  $\alpha$ -subunits might enhance their stability and remove the need for a stabilizing small oxygenase subunit.<sup>63</sup>

The oxygenase  $\alpha$ -subunit contains two metal centers, a Rieske-type [2Fe-2S] cluster that mediates electron transfer to the mononuclear iron in the active site where  $O_2$  binding and activation is thought to take place.<sup>100</sup> The active site iron is highly accessible with a large part of its surface available for the binding of both oxygen atoms.<sup>101</sup> All RO oxygenase structures determined to date possess either an  $\alpha_3$  or  $\alpha_3\beta_3$  quaternary structure (Figure 2A). In this complex, electron transfer proceeds between the Rieske cluster (blue) and the iron center (green) of adjacent subunits *via* a bridging aspartate (black, Figure 2B) due to the smaller distance compared to intramolecular transfer.<sup>63</sup>



**Figure 2:** (A) Quaternary structure of NDO oxygenase  $\alpha$ -subunits from *Pseudomonas* sp. NCIB 9816-4 (PDB code: 1NDO)<sup>102</sup> forming a trimer (single  $\alpha$ -subunits shown in green, blue and grey with the non-heme iron center shown as orange spheres and the Rieske cluster as orange-yellow spheres). (B) Electron transfer from the Rieske center (blue) to the catalytic mononuclear iron (green) of the adjacent subunit proceeds *via* a conserved aspartate residue (black), adapted from Barry and Challis.<sup>62</sup>

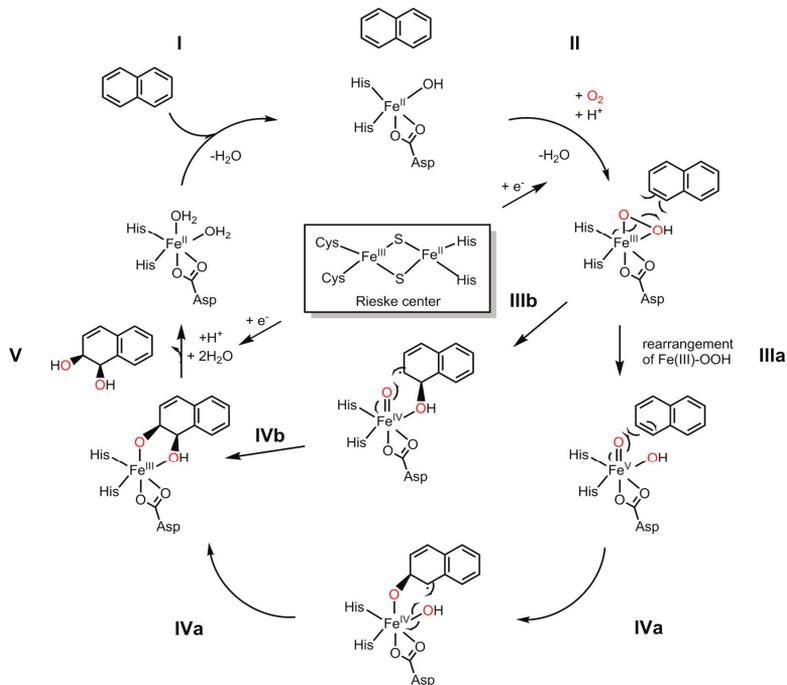
NDO from *Pseudomonas* sp. NCIB 9816-4 was the first RO for which the crystal structure has been solved and its active site geometry is characteristic for ROs. In NDO, the non-heme iron in the hydrophobic active site is coordinated by two histidines (H208 and H213) and a bidentate aspartate (D362) which constitute the 2-His-1-carboxylate facial triad.<sup>102</sup> Additionally, one or two water molecules are found to be bound in the resting state when no substrate is present (Scheme 12).<sup>103</sup> The Rieske center is coordinated by two cysteine residues (C81 and C101) and two histidines (H83 and H104) which are located in the highly conserved motif CXHX<sub>17</sub>CX<sub>2</sub>H.<sup>76,102</sup> In the  $\alpha_3\beta_3$  hexamer, the Rieske cluster of one  $\alpha$ -subunit is closer to the catalytic iron of the adjacent subunit ( $\sim 12$  Å) than to the Rieske cluster within the same subunit ( $\sim 44$  Å).<sup>103</sup> The neighboring non-heme iron

and [2Fe-2S] centers are thought to be coupled by D205 that is within hydrogen-bonding distance of the metal-coordinating amino acids and bridges H104 (coordinating the Rieske center) to H208 of the mononuclear iron in the active site. D205 is most likely the main electron transfer pathway from the Rieske cluster to the iron and replacement of this residue (NDO variant D205A) completely abolished activity.<sup>104</sup>

In contrast to cytochrome P450 monooxygenases, far less is known about the catalytic mechanism of ROs to date.<sup>62</sup> Due to difficulties in obtaining spectroscopic information of the mononuclear non-heme iron in the presence of the Rieske cluster, the mononuclear ferrous site is more challenging to access spectroscopically.<sup>103</sup> For dihydroxylation reactions, two possible mechanisms have been suggested for NDO which both are initiated by binding of naphthalene to the active site (Scheme 12).<sup>62</sup> Substrate binding converts the iron coordination geometry from a distorted octahedral geometry to a five-coordinate square pyramid due to the elimination of the water ligand (**I**, Scheme 12). The geometric and electronic structures of the catalytic site in the presence of the substrate are significantly affected by the redox state of the Rieske center. Reduction of the Rieske cluster transforms the square pyramidal geometry into a mixture of trigonal bipyramidal and square pyramidal sites, priming the catalytic iron for the activation of molecular oxygen. These protein dynamics regulate the reactivity toward O<sub>2</sub> during the catalytic cycle with both substrate binding and reduction of the [2Fe-2S] center being required for O<sub>2</sub> activation. The structural changes prevent oxygen activation in the absence of bound substrate and hence the formation of reactive oxygen species which might lead to enzyme inactivation.<sup>103,105</sup> The crystal structure of NDO from *Pseudomonas* sp. NCIB 9816-4 co-crystallized with O<sub>2</sub> (PDB code: 1O7N) shows side-on binding of oxygen to the catalytic non-heme iron center.<sup>101</sup> Coordination of O<sub>2</sub> to the iron complex in a side-on fashion is a unique feature of these enzymes and allows each oxygen atom to attack adjacent carbons from the same face of the substrate, leading to *cis*-diol formation. Other RO crystal structures show that oxygen and substrate bind in a similar manner with the atoms closest to the Fe-O<sub>2</sub> complex being oxidized. Thereby, interactions between the substrate and residues constituting the active site are assumed to orient the substrate in the enzyme binding pocket.<sup>63</sup>

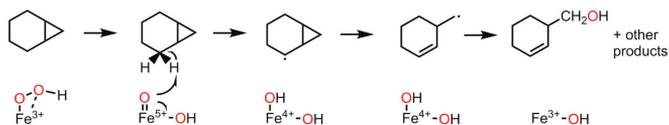
After O<sub>2</sub> binding, subsequent electron transfer from the Rieske cluster to the non-heme iron center results in a ferric peroxide complex which is transferred by subsequent protonation and loss of the water ligand into a bidentate hydroperoxide complex (Fe(III)-OOH; **II**, Scheme 12).<sup>62</sup> The formation of a (hydro)peroxo species as pathway intermediate is supported by the uncoupling of O<sub>2</sub> reduction and dihydroxylation.<sup>100</sup> Upon exposure to certain substrate analogues (e.g. benzene for NDO), uncoupled NADH and substrate oxidation generates H<sub>2</sub>O<sub>2</sub> in stoichiometric amounts, suggesting that “poor” substrates lead to a longer lifetime for an iron-peroxo species which finally decays into H<sub>2</sub>O<sub>2</sub>.<sup>21</sup> The intermediates of the next step have not been completely solved yet and it is still unknown whether a high-valent Fe(V) or Fe(IV) oxo intermediate is formed. In case O-O bond cleavage precedes substrate oxidation, a Fe(V)=O(OH) complex is the proposed intermediate (**IIIa**, Scheme 12) which subsequently reacts with the substrate (**IVa**, Scheme 12). When O-O bond cleavage and substrate oxidation are coupled, a Fe(IV)=O complex is formed in combination with a substrate radical intermediate (**IIIb**, Scheme 12) that reacts further (**IVb**, Scheme 12). Regeneration of the enzyme

resting state affords transfer of a second electron from the [2Fe-2S] cluster to the iron and protonation of the intermediate which results in the release of the oxyfunctionalized product (V, Scheme 12).<sup>62</sup>



**Scheme 12:** Proposed mechanism for the RO-catalyzed dihydroxylation, adapted from Barry and Challis.<sup>62</sup>

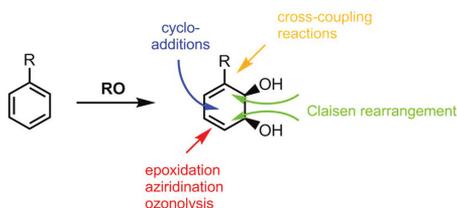
As mentioned in the previous section, ROs are able to catalyze both dihydroxylation and monohydroxylation reactions. For the introduction of a single hydroxyl group, Chakrabarty and coworkers proposed a mechanism involving radical intermediates (Scheme 13).<sup>106</sup> By using cyclopropane-containing substrates like norcaradiene and bicyclohexane as probes, formation of radical or cationic intermediates could be distinguished as either rearranged or ring expanded products are obtained. Results for conversion of norcaradiene and bicyclohexane using NDO as biocatalyst support a radical rebound mechanism for which the radical lifetime can be estimated by the ratio of unarranged to rearranged products. As already proposed for dihydroxylation reactions, the mechanism of monohydroxylation might include a high-valent Fe(V)-oxo-hydroxo intermediate (Scheme 13) and results for the oxidation of cyclopropane-containing substrate probes indicate the formation of a Fe(V) species prior to the insertion reaction.<sup>106</sup>



**Scheme 13:** Proposed mechanism for the monohydroxylation of norcaradiene by NDO.<sup>106</sup>

### 1.3.4. Applications of ROs

To date, the stereoselective *cis*-dihydroxylation of aromatic compounds has been the basis for most applications using RO enzymes. As no chemical single-stage alternative with similar broad substrate specificity and excellent stereoselectivity is available, a strong interest in ROs as biocatalysts has been developed.<sup>51</sup> Due to their unique ability to generate arene *cis*-dihydrodiols, ROs are used for the preparation of lead structures found in natural products, polyfunctionalized metabolites and pharmaceutical intermediates.<sup>82</sup> Arene-1,2-dihydrodiols are versatile chiral building blocks that can be further functionalized in various ways (Scheme 14),<sup>107</sup> e.g. by replacing the hydroxyls with other functional groups or using the two chiral centers to direct stereo-controlled Diels-Alder reactions at or Michael addition reactions to the cyclohexadiene double bonds.<sup>21</sup>



**Scheme 14:** Functionalization of arene *cis*-dihydrodiols by cycloaddition, cross-coupling reactions, Claisen rearrangement, epoxidation, aziridination or ozonolysis, adapted from Schrittwieser and Resch.<sup>107</sup>

Within the last decades, there has been an enormous increase in the application of RO enzymes in organic synthesis.<sup>3,108</sup> In particular the research group of Hudlicky has focused on the application of enzymatically derived arene *cis*-dihydrodiol units as chiral building blocks in synthetic organic chemistry. Starting from cyclohexadiene-1,2-diols as intermediates, various compounds including inositols, conduritols and cyclitols,<sup>108</sup> strawberry furanone,<sup>109</sup> prostaglandin E<sub>2</sub>,<sup>110</sup> morphine<sup>111</sup> and pancratistatin<sup>111</sup> were synthesized in different chemoenzymatic approaches. To overcome the problem that the highly stereoselective RO-catalyzed dihydroxylation only provides access to one enantiomer, Hudlicky and coworkers employed the TDO-mediated oxidation of *para*-dihalobenzenes with a iodine substituent that results in the opposite configuration of the products compared to the unsubstituted compounds, followed by removal of the iodine atom.<sup>112</sup>

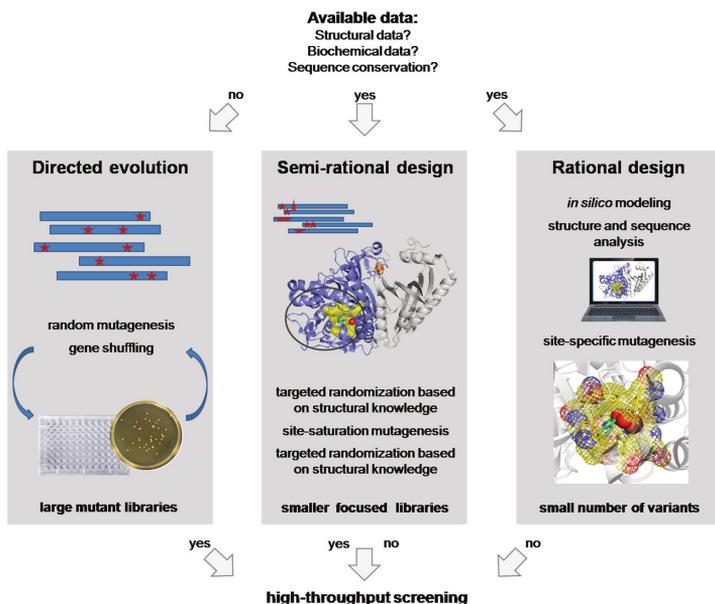
To date, several industrial applications of RO-catalyzed *cis*-dihydroxylations have been described. In 1983, the Imperial Chemical Industries group reported a biotechnological route to polyphenylene with *meso-cis*-cyclohexanedienediol as intermediate.<sup>113</sup> Another application is the RO-catalyzed formation of the blue dye indigo.<sup>114</sup> The biotechnological production of indigo starts from glucose with metabolically engineered *Escherichia coli* (*E. coli*) strains expressing NDO.<sup>21</sup> In the process, NDO catalyzes the dihydroxylation of indole formed in the tryptophan biosynthetic pathway and the resulting *cis*-diol undergoes spontaneous dehydration and dimerization to indigo. Furthermore, Merck applied ROs for the biocatalytic production of *cis*-aminoindan-2-ol, a key intermediate in the synthesis of the HIV inhibitor Crixivan®. The process relies on the biotransformation of indene to *cis*-1,2-indandiol and TDO forms the desired (1*S*,2*R*)-enantiomer which can be used as synthon for the preparation of (1*S*,2*R*)-amino-indan-2-ol. However, the process is hampered due to by-product formation (1-indenol)

and low stereoselectivities that can be overcome by the presence of a dihydrodiol dehydrogenase in *P. putida* (*ee* > 98%). As alternative strategy, a *Rhodococcus* strain harboring a suitable RO system was employed for the formation of *cis*-1,2-indandiol with good stereoselectivities.<sup>9</sup>

Next to the generation of chiral building blocks, ROs are applied as a platform for bioremediation. Due to its broad substrate scope regarding aromatic compounds, biphenyl dioxygenase (BphDO) has been extensively studied for the biodegradation of arenes with an emphasis on polychlorinated biphenyls (PCBs) which are serious environmental pollutants.<sup>63</sup> Next to PCBs, carcinogenic chloroethenes like trichloroethylene (TCE) found in soil and groundwater<sup>115</sup> have been shown to be transformed by TDO into formate and glyoxylate.<sup>82</sup>

#### 1.4. Protein engineering – Different strategies for the tailoring of ROs

Even though the application of enzymes in organic synthesis offers many advantages due to their intrinsic high chemo-, regio- and stereoselectivity,<sup>52</sup> potential biocatalytic routes might not always be efficient enough compared to chemical synthesis.<sup>116,117</sup> Protein engineering allows improving the performance of enzymes as catalysts in synthetic organic chemistry and biotechnology. Two main strategies, rational design and directed evolution can be employed to alter enzyme properties for the desired application (Figure 3).<sup>118</sup> Both strategies or a combination thereof (semi-rational design) have been described for the engineering of ROs to alter their substrate scope, activity or selectivity.<sup>119</sup> The best studied examples include NDO, TDO and BphDO which have been subjected to various engineering methods in order to identify residues controlling selectivity and activity.<sup>82</sup>



**Figure 3:** Protein engineering approaches including directed evolution, rational design and semi-rational design, adapted from Steiner and Schwab.<sup>120</sup>

### 1.4.1. Rational enzyme design

In rational enzyme design, site-specific mutagenesis is applied to introduce amino acid changes in order to alter or induce desired properties.<sup>121</sup> To be able to predict the most promising sites for mutagenesis in the protein sequence, detailed knowledge about structural and mechanistic features is required and often combined with *in silico* modeling.<sup>116</sup> Yet, the required structural and mechanistic data are available for only a small number of proteins. Furthermore, one of the most significant limitations in the general application of rational design is the complexity of the structure-function relationship in enzymes and predicting the effect of mutations on protein features proved to be difficult.<sup>119,122</sup>

Despite these limitations, rational design was successfully applied for the mutagenesis of different ROs, mainly targeting residues located in the substrate binding pocket. For many biocatalysts, the majority of mutations that beneficially affect enzyme properties like enantioselectivity or substrate specificity are rather located in or near the active site and might include amino acid residues that are involved in substrate binding.<sup>123</sup> Several residues in the active site of NDO were identified by site-directed mutagenesis resulting in altered selectivity and product formation compared to the wild type enzyme.<sup>124–126</sup> In particular, position F352 in NDO from *Pseudomonas* sp. strain NCIB 9816-4 plays a major role in controlling both the regio- and stereoselectivity towards biphenyl and phenanthrene, offering the possibility to engineer dioxygenases by the introduction of a single point mutation.<sup>124</sup> The corresponding position, F366 in tetrachlorobenzene dioxygenase from *Ralstonia* sp. PS12 was also shown to influence the regioselectivity of the enzyme. By remodeling of the active site using site-directed mutagenesis, variant F366L displayed significant changes in regioselectivity for 2-, 3-chloro-, 2,4-, 2,5- and 2,6-dichlorotoluene.<sup>127</sup> Furthermore, an alteration in regiospecificity of BphDO from *Pseudomonas pseudoalcaligenes* KF707 for biphenyl-related compounds could be achieved by active site engineering.<sup>128</sup> Variants F227V, I335F, T376N, F377L and F377A showed altered regioselectivities for dihydroxylation of various PCBs compared to the wild type enzyme. In addition, the change in four amino acids (positions 335, 336, 338 and 341) by site-directed mutagenesis of BphDO from *Pseudomonas* sp. LB400 enhanced its capacity to degrade PCBs by extending the substrate range of BphDO.<sup>129</sup>

Yet, rational site-directed approaches are restricted due to the limited sequence space that can be explored at a time. In contrast, directed evolution of enzymes offers the possibility to randomly introduce mutations to cover a large sequence space. Using the random approach, also mutations far from the active site that influence distinct enzyme properties like activity or thermostability can be identified.<sup>123</sup>

### 1.4.2. Directed evolution

In directed evolution, random changes are introduced into a gene of interest, altering the amino acid sequence of an enzyme in an iterative manner until the desired properties are obtained.<sup>117</sup> It is the most widely applied approach for protein engineering as new functional properties can be evolved in enzymes without prior knowledge about the structure or function of the target protein.<sup>122,130</sup> Directed

evolution is based on two iterative steps which involve repeating cycles of mutagenesis for diversity generation and screening or selection for the identification of improved variants.<sup>131</sup>

In directed evolution experiments, molecular diversity is created either by *in vitro* gene recombination methods or by random mutagenesis.<sup>131</sup> During recombination, a set of related sequences is randomly combined whereas non-recombining methods only target a single protein sequence.<sup>116</sup> Among the most often used mutagenesis methods in directed evolution are error-prone PCR (epPCR), saturation mutagenesis and shuffling approaches.<sup>121</sup> Often, a combination of both random mutagenesis and recombination is applied as deleterious mutations introduced by methods like epPCR can be crossed out by shuffling and only beneficial mutations are preserved.<sup>132</sup> By directed evolution, a large sequence space can be covered, generating a huge number of mutants that must be screened. Hence, a reliable high-throughput screening (HTS) or selection assay is needed to identify functionally improved variants under relevant conditions. The development of a suitable HTS assay is one of the major limitations of directed evolution and the success of random mutagenesis experiments often depends on the availability of screening or selection methods.<sup>122,131</sup>

Directed evolution of ROs has helped to identify residues that effect enzyme function.<sup>63</sup> A widely applied method for the introduction of mutations into ROs was the use of *in vitro* gene recombination methods like DNA shuffling, staggered extension process (StEP) or random priming recombination.<sup>21,133,134</sup> These methods mimic nature by random recombination of selected genes *in vitro* and thus offer an attractive approach as RO enzymes can be grouped into families similar in size and amino acid sequence.<sup>64</sup> Herein, enzymes belonging to the same family have evolved from a common ancestor to acquire new catalytic functions through recombination, duplication, multiple point mutations, deletion or integration, resulting in diverse but highly related sequences<sup>135</sup> and studies of hybrid ROs have shown altered selectivities and product formations.<sup>63</sup> Using recombination methods, one extensively studied example in directed evolution experiments is BphDO that has been evolved for bioremediation applications by extending its substrate scope. For BphDOs from different aromatics degrading strains, variations in the PCB congener selectivity patterns and activities were found.<sup>135</sup> BphDOs from *Pseudomonas pseudoalcaligenes* strain KF707 and *Pseudomonas cepacia* LB400 differ in their substrate range despite their nearly identical amino acid sequence of the oxygenase  $\alpha$ -subunits which vary in only 20 amino acid residues (96.4% identity). By construction of chimeric BphDOs from both strains, a small number of amino acid residues at the C-terminus was identified that are involved in substrate specificity and regioselectivity, being responsible for PCB recognition. Using three restriction sites common to the  $\alpha$ -subunits of BphDO oxygenases from both strains, a KF707 variant with an expanded range of biodegradable PCB congeners could be generated.<sup>136</sup> Based on these results, DNA shuffling of *bphA1* genes encoding for the BphDO  $\alpha$ -subunits from *Pseudomonas pseudoalcaligenes* KF707 and *Pseudomonas cepacia* LB400 was used to recombine as many differences in both enzymes as possible.<sup>135</sup> Several evolved BphDO variants could be identified based on the formation of yellow colored *meta*-cleavage products, showing enhanced degradation capacity for PCB and related biphenyl compounds. In addition, novel degradation capacity for smaller aromatic substrates like benzene and toluene was detected which are only poorly attacked by wild type BphDOs.<sup>135</sup>

Yet, in these DNA shuffling experiments by Kumamaru and coworkers, no chimeric enzymes were found that combine the ability of both parental clones to dihydroxylate double *ortho*- and *para*-substituted PCBs. Whereas strain LB400 preferentially dihydroxylates *ortho*-substituted PCBs, *para*-substituted PCBs are oxidized by *Pseudomonas pseudoalcaligenes* KF707. Applying DNA shuffling and StEP to recombine BphDO  $\alpha$ -subunits, Brühlman and Chen identified several variants that equally degraded both PCB substitution patterns and also showed enhanced dihydroxylation activity for several tetra- and pentachlorinated PCBs being hardly degraded by wild type BphDOs.<sup>133</sup> Interestingly, both Kumamaru *et al.* and Brühlman and Chen identified variants T335A and F336I during their mutagenesis studies, indicating a crucial effect of these positions on selectivity.<sup>133,135</sup>

Besides the various recombinative approaches for RO engineering, also examples for the use of non-recombinative evolutionary methods exist. Random mutagenesis was applied to increase the activity of aniline dioxygenase for aniline derivatives. In order to enhance the bioremediation of aromatic amines, the mutant V205A was used as template for one round of saturation mutagenesis on the active site residues followed by a round of epPCR. Even though variant V205A exhibited an extended substrate specificity for 2-isopropylaniline compared to the wild type enzyme, replacement of valine by alanine at position 205 decreased the activity for aniline (8.4 x) and 2,4-dimethylaniline (28 x). Applying random mutagenesis, aniline dioxygenase activity was improved, yielding several mutants with increased dihydroxylation capacity.<sup>137</sup> Likewise, in order to extend the substrate scope of TDO towards 4-picoline, a combination of epPCR and saturation mutagenesis was applied and screening of mutant libraries was accomplished using Gibbs' reagent (2,6-dichloroquinone-4-chloroimide)<sup>138</sup> which strongly reacts with *ortho*- and *meta*-substituted phenolic compounds to form colored products.<sup>139</sup>

#### 1.4.3. Semi-rational design: Combining directed evolution and rational design

As mentioned in the previous sections, both rational enzyme design and directed evolution have distinct limitations. Whereas rational design is restricted due to the insufficient knowledge on protein-structure relationships, the availability of fast and reliable HTS methods for screening of large library sizes can be limiting in directed evolution experiments.<sup>116</sup> In addition, even in huge protein libraries, only a fraction of the possible protein variants is represented and biases resulting from the degeneracy of the genetic code further restrict diversity.<sup>140</sup> Taking these problems into account, the focus has moved onto new strategies for the design of smaller libraries with high quality that eliminate the need for suitable HTS assays. For library design, a combination of rational design and directed evolution approaches is applied to overcome the limitations of each method and combine the advantages of both strategies to create small focused mutant libraries.<sup>116,118</sup>

In semi-rational design, the choice of positions to mutate remains rational in most cases and the reduced library size results from the focus on specific amino acid residues that are targeted on the basis of prior structural or functional knowledge. Sequence-based enzyme engineering uses multiple sequence alignments and phylogenetic analyses to determine amino acid conservation and evolutionary relationships among groups of homologous proteins.<sup>141</sup> Furthermore, structural analysis can facilitate the mutation of enzymes in a more targeted manner and due to the growing number of structures and advances in homology modeling, structure-based enzyme redesign allows the effective location of key residues.<sup>141</sup> The resulting focused mutant libraries are more likely to yield positive

results as multiple mutations can be introduced at positions where they might be most effective e.g. in or near the active site when targeting enzyme properties like enantioselectivity, substrate specificity or new catalytic activities.<sup>122,123</sup>

Site-directed saturation mutagenesis was applied to improve the activity and selectivity of nitrobenzene dioxygenase from *Comamonas* sp. for the selective oxidation of aromatic sulfides. Targeting different residues in the active site generated dioxygenase variants with opposite enantioselectivities and increased activities on substrates like thioanisole, Cl- or Br-thioanisole and *p*-tolyl compared to the wild type enzyme.<sup>142</sup> Furthermore, the BphDO-catalyzed dihydroxylation of 7-hydroxyflavone and 5,7-dihydroxyflavone was successfully enhanced by site-saturation mutagenesis at positions 324 and 325 of the best BphDO variants obtained from family shuffling of the oxygenase  $\alpha$ -subunits from *Pseudomonas pseudoalcaligenes* strain KF707 and *Pseudomonas putida* strain KF715.<sup>143</sup>

## 1.5. Aim of the work

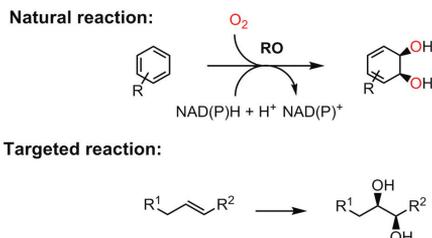
Due to the increasingly stringent environmental constraints and with regard to atom efficiency, novel catalytic methodologies for organic synthesis are in great demand by the pharmaceutical and chemical industries.<sup>144</sup> Being part of the project “Chemistry for the 21<sup>st</sup> century” (CHEM21) with the aim to establish sustainable alternatives to existing chemical methods, this PhD project has focused on a biocatalytic access to *cis*-1,2-diols. Employing enzymatic routes for alkene *cis*-dihydroxylation would reduce the need for toxic heavy metal catalyst which can lead to byproduct formation due to over-oxidation and ring fission.<sup>5</sup> In order to provide efficient biocatalysts for the formation of vicinal *cis*-diols, the following objectives were targeted:

- Identification of suitable biocatalysts for the asymmetric dihydroxylation of alkenes
- Tailoring these enzymes for the oxyfunctionalization of a selected panel of olefins
- Characterization of biocatalysts and comparison with chemical means
- Proof of the applicability of designed biocatalysts for organic synthesis

In a first step, the identification and selection of enzymes suitable for the oxyfunctionalization of olefins should be performed. As ROs have been described to be the only enzymes capable of inserting both atoms of O<sub>2</sub> in one enzymatic step to stereoselectively catalyze the formation of vicinal *cis*-diols,<sup>63</sup> this enzyme class was further investigated. Being naturally involved in the dihydroxylation of aromatic compounds,<sup>65</sup> RO enzymes had to be selected that show an extended substrate specificity towards unnatural alkene substrates. These enzymes were to be characterized regarding their substrate scope and selectivities using a panel of structurally different alkenes to give a first overview of the olefins that can be targeted employing ROs. To evaluate biocatalytic olefin oxidation against heavy metal catalyzed methods like Sharpless AD, the panel of compounds should also include alkenes difficult to target with OsO<sub>4</sub> as catalyst.

In the second step, wild type enzymes were to be optimized by protein engineering for the selective oxyfunctionalization of the alkene substrates. Rational and semi-rational design should be applied as mutagenesis strategies to create a small focused mutant library. Within this mutant library, amino acid residues crucial for selectivity and substrate specificity of ROs should be identified. Generated variants were to be characterized and compared to both wild type enzymes and chemical methods for olefin dihydroxylation.

In the third step, the applicability of RO-catalyzed olefin oxidation in organic synthesis should be demonstrated by performing semi-preparative biotransformations in mg scale.



**Scheme 15:** Natural reaction of ROs and the dihydroxylation of alkenes addressed in this work.

## 2. EXPERIMENTAL SECTION

### 2.1. Materials

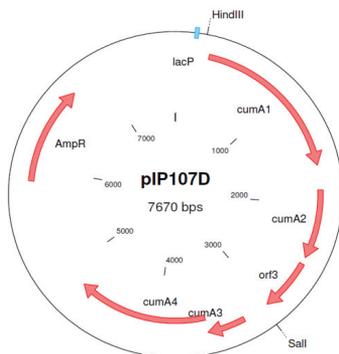
Except not noted otherwise, chemicals, solvents and buffer components were obtained from Sigma-Aldrich and Fluka (Steinheim, Germany), Carl Roth GmbH (Karlsruhe, Germany), Acros Organics (Geel, Belgium), VWR (Darmstadt, Germany), Alfa Aesar (Karlsruhe, Germany) and Serva (Heidelberg, Germany). Restriction enzymes as well as the T4 DNA ligase were purchased from Thermo Scientific (Schwerte, Germany), *Pfu* polymerase from GeneOn (Ludwigshafen, Germany) and enzymes for Gibson Assembly from NEB (Frankfurt am Main, Germany). Kits for PCR product purification, gel extraction of DNA and plasmid minipreparations were obtained from Zymo Research (Freiburg, Germany). DNA markers for agarose gel electrophoresis as well as protein markers for SDS-PAGE were obtained from peqlab (peqlab Biotechnology GmbH, Erlangen, Germany) or Fermentas (Thermo Fisher Scientific Germany, Braunschweig). peqGREEN for DNA staining in agarose gel electrophoresis was purchased from peqlab (peqlab Biotechnology GmbH, Erlangen, Germany). DNA sequencing was performed by GATC Biotech (Konstanz, Germany) and all primers were synthesized by metabion International AG (Martinsried, Germany).

### 2.2. Plasmids, strains and storage of *E. coli* cultures

The plasmid pDTG141 carrying the *nah* genes encoding for NDO oxygenase  $\alpha$ - (*nahAc*) and  $\beta$ - (*nahAd*) subunits, ferredoxin (*nahAb*) and reductase (*nahAa*) from *Pseudomonas* sp. strain NCIB 9816-4 was obtained from Rebecca Parales.<sup>145</sup> The plasmid pJRM501 containing the genes for BDO oxygenase  $\alpha$ - (*bedC1*) and  $\beta$ - (*bedC2*) subunits, ferredoxin (*bedB*) and reductase (*bedA*) from *P. putida* ML2 was provided by David Leak.<sup>146</sup> The plasmid pIP107D harboring the genes for CDO oxygenase  $\alpha$ - (*cumA1*) and  $\beta$ - (*cumA2*) subunits, ferredoxin (*cumA3*) and reductase (*cumA4*) from *Pseudomonas fluorescens* IP01 was obtained from Hideaki Nojiri.<sup>147</sup>



**Figure 4:** The plasmids pJRM501 (left) and pDTG141 (right) encoding the RO genes for BDO and NDO. pJRM501 is based on the pKK223-3 vector containing an ampicillin resistance gene (AmpR) and BDO genes are under the control of the IPTG-inducible *tacP*.<sup>148</sup> pDTG141 is based on the pT7-5 vector harboring an ampicillin resistance gene (AmpR) and NDO genes are under the control of the T7 promoter (T7P).<sup>92</sup>



**Figure 5:** Plasmid pIP107D encoding the RO genes for CDO. pIP107D is based on the pUC118 vector containing an ampicillin resistance (AmpR) and genes are under the control of the IPTG-inducible *lacP* promoter (lacP).<sup>147</sup>

For cloning, *E. coli* XL1-Blue was employed, a strain deficient in endonuclease (*endA*) and recombination (*recA*). The cleavage of cloned DNA by the *EcoK* endonuclease system is prevented by the *hsdR* mutation (Table 4). For expression of RO genes, *E. coli* strains JM109 (CDO and BDO) and JM109 (DE3) (NDO) were used. *E. coli* JM109 (DE3) carries a chromosomal copy of the T7 RNA polymerase gene under the control of the *lacUV5* promoter (Table 4).

**Table 4:** Different *E. coli* strains used for cloning and expression of RO genes.

Strain	Genotype
<i>E. coli</i> XL1-Blue	<i>recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac</i> [F' <i>proAB lac<sup>q</sup>ΔM15 Tn10</i> (Tet <sup>r</sup> )]
<i>E. coli</i> JM109	<i>endA1 recA1 gyrA96 thi hsdR17 (rk<sup>-</sup> mk<sup>+</sup>) relA1 supE44</i> $\Delta$ ( <i>lac-proAB</i> ) [F' <i>traD36 proAB laq<sup>q</sup>ΔM15</i> ]
<i>E. coli</i> JM109 (DE3)	<i>endA1 recA1 gyrA96 thi hsdR17 (rk<sup>-</sup> mk<sup>+</sup>) relA1 supE44 λ-</i> $\Delta$ ( <i>lac-proAB</i> ) [F' <i>traD36 proAB lac<sup>q</sup>ΔM15</i> ] <i>IDE3</i>

For long term storage of *E. coli* strains at -80°C, glycerol stocks were prepared by the addition of sterile glycerol to *E. coli* overnight cultures grown in lysogeny broth (LB) medium to a final concentration of 15% (v/v).

## 2.3. Methods

### 2.3.1. Restriction enzyme-based gene cloning - Reconstruction of plasmid pIP107D

The plasmid pIP107D that was obtained from Hideaki Nojiri lacked part of the reductase gene *cumA4*. Thus, the reductase was restored using plasmid pIP103 (provided by Hideaki Nojiri) as template for PCR which harbored the corresponding *cis*-dihydrodiol dehydrogenase and the extradiol dioxygenase next to the RO components.<sup>147</sup> By touchdown PCR,<sup>149</sup> an *EcoRI* restriction site was appended after the stop codon of *cumA4* using the primers given in Table 5. The plasmid pIP107D as well as the PCR product were digested with *SalI* and *EcoRI* restriction enzymes according to the manufacturer's instructions. Digested DNA was purified with the *Zymoclean™ Gel DNA Recovery Kit* (plasmid DNA)

or with the *DNA Clean & Concentrator-5™* kit (PCR product) and ligated in a ratio of 3:1 (PCR product:plasmid DNA) with T4 DNA ligase over night at room temperature according to the manufacturer's instructions. After heat inactivation of the T4 DNA ligase, 2.5 µl of the ligation mixture were transformed in 50 µl chemically competent *E. coli* XL1-Blue cells and plated on LB agar (1.5%) supplemented with 100 µg mL<sup>-1</sup> ampicillin. Colonies were picked and grown over night for minipreparations. The correct sequence was confirmed by DNA sequencing (primers in Table S24).

**Table 5:** Primers used to amplify the *cumA4* gene and append an *EcoRI* restriction site (underlined) after the stop codon. fwd: forward primer, rev: reverse primer.

Primer name	Sequence
pIP107D 4_fwd	5'-GCC AATCTCCTCGGGACTTTGC-3'
pIP107D cumA4 <i>EcoRI</i> _rev	5'-CATAATGAATTCTCACTCGCATCGCTCAGCTTTAG-3'

### 2.3.2. Site-directed mutagenesis

Site-directed mutagenesis of CDO (CDO M232A) and NDO (NDO A206I and NDO A206M) was performed according to the Stratagene QuikChange™ protocol<sup>150</sup> using the primers given in Table 6. To introduce mutation M220A in BDO, one step isothermal Gibson assembly<sup>151</sup> was applied as QuikChange™ PCR did not give the desired results. Plasmid pJRM501 carrying the *bed* genes was digested with *EcoRI* and vector DNA gel purified with the *Zymoclean™ Gel DNA Recovery Kit*. To create overlapping parts with the digested plasmid, BDO genes were amplified using standard PCR protocols with the given primers (BDO M220A\_fwd and GA BDO M220A\_rev/ BDO M220A\_rev and GA BDO M220A\_fwd, Table 6) and the PCR reactions were digested for 1 h at 37°C with *DpnI* to remove the template DNA. After Gibson assembly<sup>151</sup> of the plasmid DNA and the PCR product, 2.5 µl of the reaction mixture were transformed in 50 µl chemically competent *E. coli* XL1-Blue cells and plated on LB agar (1.5%) supplemented with 100 µg mL<sup>-1</sup> ampicillin. The correct DNA assembly and the insertion of point mutations was confirmed by DNA sequencing (primers given in Table S24) after minipreparation of plasmids from overnight cultures of grown colonies.

**Table 6:** Primers used for site-directed mutagenesis of BDO, NDO and CDO and the generation of overlapping PCR products for Gibson assembly of BDO. Nucleotides to insert point mutations are underlined. fwd: forward primer, rev: reverse primer.

Primer name	Sequence
BDO M220A_fwd	5'-GTTTTGTAGTGAC <u>CCG</u> TACCATGCCGGGAC-3'
BDO M220A_rev	5'-GTCCCGGCATGGTAC <u>CCG</u> CTCACTACAAAAC-3'
GA BDO M220A_fwd	5'-CGACATCATAACGGTTCCTGGC-3'
GA BDO M220A_rev	5'-GAGTTCGGCATGGGGTCAG-3'
CDO M232A_fwd	5'-CTGTAGCGAT <u>CCG</u> TACCATGCCGGGAAC-3'
CDO M232A_rev	5'-GTTCCCGCATGGTAC <u>CCG</u> ATCGCTACAG-3'
NDO A206I_fwd	5'-CTTTGTGGGAGAT <u>ATCT</u> ACCACGTGGGTTG-3'
NDO A206I_rev	5'-CAACCCACGTGGTAG <u>AT</u> ATCTCCCAAAAG-3'
NDO A206M_fwd	5'-CTTTGTGGGAGAT <u>ATGT</u> ACCACGTGGGTTG-3'
NDO A206M_rev	5'-CAACCCACGTGGTAC <u>AT</u> ATCTCCCAAAAG-3'

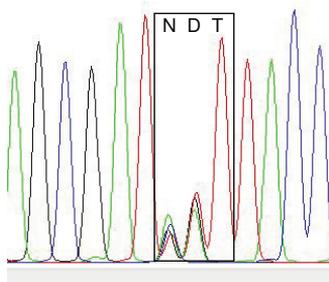
### 2.3.3. Site-saturation mutagenesis and indole screening assay

For site-saturation mutagenesis of CDO at positions F278, I288, I336 and F378, primers with DBC or BBT degenerated codons were used encoding for 8 (DBC: F, C, I, T, S, V, A, G) and 9 (BBT: F, S, C, L, P, R, V, A, G) mainly small amino acids. For mutagenesis at position M232, the NDT degenerated codon was used to introduce a balanced mix of 12 polar and nonpolar, aliphatic and aromatic as well as negatively and positively charged amino acids (F, L, I, V, Y, H, N, D, C, S, R and G).<sup>140</sup> DBC, BBT and NDT libraries were prepared using the primers with the degenerated codons given in Table 7. For BBT and DBC mutant libraries, overlapping DNA fragments containing the introduced mutations (forward primers for introduction of point mutations and pIP107D cumA3\_rev/ reverse primers for introduction of point mutations and pIP107D PP\_fwd) were generated using standard PCR protocols. The PCR reactions were digested for 1 h at 37°C with *DpnI* to remove the template DNA. Plasmid pIP107D was digested with *HindIII* and *SalI* and DNA gel purified with the *Zymoclean™ Gel DNA Recovery Kit*. PCR products and digested vector DNA were assembled using one step isothermal DNA assembly after Gibson *et al.*<sup>151</sup> to generate the circular vector containing the mutation. For the NDT mutant library, the QuikChange™ mutagenesis method according to the Stratagene protocol<sup>150</sup> was applied and one step isothermal Gibson assembly<sup>151</sup> was performed to seal nicks. Chemically competent *E. coli* JM109 cells were transformed with 2.5 µl of the Gibson assembly mixtures and selected on LB agar containing 100 µg mL<sup>-1</sup> ampicillin. For a coverage of ≥ 95%, at least 25 (BBT and DBC libraries) and 65 (NDT library) colonies, respectively, were screened<sup>b</sup> and a balanced library was confirmed by sequencing the resuspended clones of the whole plate (example shown in Figure 6). DNA sequencing confirmed the correct DNA assembly and insertion of point mutations (primers given in Table S24).

**Table 7:** Primers applied for the site-saturation mutagenesis of CDO at positions M232, F278, I288, I336 and F378. Degenerated codons are underlined. fwd: forward primer, rev: reverse primer.

Primer name	Sequence
CDO M232NDT_fwd	5'-CTGTAGCGAT <u>NDI</u> TACCATGCGGGAAC-3'
CDO M232NDT_rev	5'-GTTCCCGCATGGTA <u>AHN</u> ATCGTACAG-3'
CDO F278DBC_fwd	5'-CATGGGACCGGCTGG <u>DBC</u> AATGACGATTTCCGC-3'
CDO F278DBC_rev	5'-GCGAAATCGTCATT <u>GVH</u> CCAGCCGGTCCCATG-3'
CDO I288BBT_fwd	5'-CACTTCTGCAAGCC <u>BBT</u> ATGGGTCCTAAGGTTG-3'
CDO I288BBT_rev	5'-CAACCTTAGGACCCAT <u>AVV</u> GGCTTGAGAAGTG-3'
CDO I336BBT_fwd	5'-CATTCTTCTGCGC <u>BBT</u> AATACAGTCCGTAC-3'
CDO I336BBT_rev	5'-GTA CGG ACT GTATT <u>AVV</u> GCCAGGAAGA AATG-3'
CDO F378DBC_fwd	5'-GAAAAACATCTTACC <u>DBC</u> AATCAAGG GGGAACC-3'
CDO F378DBC_rev	5'-GGTCCCCCTTGATT <u>GVH</u> GGTGAAGATGTTTTTC-3'
pIP107D PP_fwd	5'-CTGGAAGCGGGCAGTG-3'
pIP107D cumA3_rev	5'-GGTCAAGCCTTTTTCAGATTTTC-3'

<sup>b</sup> The number of colonies that have to be screened for 95% coverage was calculated by CASTER.<sup>223</sup>



**Figure 6:** Randomization efficiency of each nucleotide (black: G, blue: C, red: T, green: A) for the NDT mutant library at positions encoding M232 in CDO (black box). N = A, G, T or C; D = A, G or T.

For screening of DBC and BBT mutant libraries, a colorimetric solid phase assay based on indigo formation was applied. *E. coli* JM109 cells transformed with the mutant libraries were grown over night at 37°C on LB agar plates supplemented with 100 µg mL<sup>-1</sup> ampicillin. After colony formation, dried Whatman filter papers soaked in a 10% solution of indole in acetone were placed into petri dish covers and incubated at room temperature until a blue color formation was observed.<sup>145</sup>

#### 2.3.4. Preparation of chemically competent *E. coli* cells and transformation of *E. coli*

The preparation of chemically competent *E. coli* cells was performed according to the rubidium chloride (RbCl) method as described in literature.<sup>152</sup>

Heat-shock transformation of chemically competent *E. coli* cells was performed at 42°C for 45 s after incubating the cells mixed with the plasmid DNA for 30 min on ice. The heat-shock treatment was followed by cooling on ice for 2 min and cells were subsequently incubated for 1 h at 37°C and 180 rpm in SOC medium (5 g L<sup>-1</sup> yeast extract, 20 g L<sup>-1</sup> tryptone, 0.584 g L<sup>-1</sup> NaCl, 0.186 g L<sup>-1</sup> KCl, 20 mM MgSO<sub>4</sub> and 20 mM glucose). After incubation, *E. coli* cells were plated on LB agar (1.5%) supplemented with 100 µg mL<sup>-1</sup> ampicillin.

#### 2.3.5. Heterologous expression of RO wild type enzymes and variants

Expression of NDO and BDO was carried out as described previously.<sup>145,153</sup> *E. coli* JM109 cells harbouring the plasmid pIP107D were grown at 37°C and 180 rpm in LB medium (10 g L<sup>-1</sup> tryptone, 5 g L<sup>-1</sup> yeast extract and 5 g L<sup>-1</sup> NaCl) containing 50 µg mL<sup>-1</sup> ampicillin. At an OD<sub>600</sub> of 0.5-0.7, the culture was supplemented with 0.2 mM IPTG to induce expression of CDO. Cells were harvested after cultivation for 2 h at 30°C and 180 rpm by centrifugation (9,000 × g, 20 min, 4°C). For expression of RO variants, conditions described for the wild type enzymes were applied. To avoid loss of activity, whole cell biocatalysts were immediately used for *in vivo* biotransformations. As expression levels of ROs were low and overexpressed protein bands hardly visible in SDS-PAGE, expression of active enzyme was routinely controlled by the conversion of indole to form the blue dye indigo.

### 2.3.6. *In vivo* biotransformations with wild type and mutant enzymes

Reaction mixtures in *in vivo* biotransformations contained 0.2 g cells mL<sup>-1</sup> (*cww*) in 100 mM potassium phosphate buffer pH 7.2, 20 mM glucose for *in situ* cofactor regeneration, 50 µg mL<sup>-1</sup> ampicillin and 10 mM substrate in a total volume of 1 mL. Due to the volatility of several of the tested substrates, reactions were performed in gas tight screwed capped glass vials ( $V = 20$  mL). Biotransformations were incubated at 30°C and 180 rpm for 24 h. Negative controls were performed with *E. coli* JM109 cells carrying the empty vector.

For GC analysis, cells were centrifuged and 500 µL supernatant was transferred in an Eppendorf tube and supplemented with NaCl and 1-decanol as internal standard, followed by three times extraction with 250 µL ethyl acetate.

Preparative biotransformations for product characterization were performed in 1 L flasks with 0.25 g cells mL<sup>-1</sup> (*cww*) and 70 mg substrate in a total volume of 55 mL, applying the same reaction set-up as described above. After cell lysis by ultrasonication and addition of NaCl, the reaction mixture was extracted four times with 50 mL ethyl acetate. The combined organic fractions were concentrated *in vacuo* and subjected to flash column chromatography.

### 2.3.7. Up-scaling of (*R*)-limonene hydroxylation to 4 L

For up-scaling of (*R*)-limonene hydroxylation in a bioreactor, the protocol described by Yildirim *et al.* for the biotransformation of benzonitrile in a 2 L reactor with *E. coli* JM101 expressing chlorobenzene dioxygenase was adapted for CDO M232A.<sup>154</sup> Biotransformations with *E. coli* JM109 expressing the CDO variant M232A were performed in a Labfors stirred tank reactor (Infors AG, Bottmingen, Switzerland) with a total volume of 4 L terrific broth (TB) media (12 g L<sup>-1</sup> tryptone, 24 g L<sup>-1</sup> yeast extract, 4 g L<sup>-1</sup> glycerol mixed 10:1 with 0.17 M KH<sub>2</sub>PO<sub>4</sub> and 0.72 M K<sub>2</sub>HPO<sub>4</sub>) supplemented with 50 µg mL<sup>-1</sup> ampicillin. The temperature was kept at 30°C and the pH was adjusted to 7.2 using 25% NH<sub>4</sub>OH and 10% H<sub>3</sub>PO<sub>4</sub>. The reactor was aerated with 2 vvm and stirred at 1000 rpm. After inoculation with an OD<sub>600</sub> of 0.05, cells were supplemented with 7 g L<sup>-1</sup> (NH<sub>4</sub>)<sub>2</sub>HPO<sub>4</sub> as additional N-source and grown overnight (~12 h). The resulting culture with a cell dry weight (*cdw*) of about 7-8 g L<sup>-1</sup> was supplemented with 15 ml trace element solution US\* (4.87 g L<sup>-1</sup> FeSO<sub>4</sub>·7 H<sub>2</sub>O, 4.12 g L<sup>-1</sup> CaCl<sub>2</sub>·H<sub>2</sub>O, 1.50 g L<sup>-1</sup> MnCl<sub>2</sub>·4 H<sub>2</sub>O, 1.50 g L<sup>-1</sup> MnCl<sub>2</sub>·4 H<sub>2</sub>O, 1.05 g L<sup>-1</sup> ZnSO<sub>4</sub>, 0.30 g L<sup>-1</sup> H<sub>3</sub>BO<sub>3</sub>, 0.25 g L<sup>-1</sup> Na<sub>2</sub>MoO<sub>4</sub>·2 H<sub>2</sub>O, 0.15 g L<sup>-1</sup> CuCl<sub>2</sub>·2 H<sub>2</sub>O and 0.15 g L<sup>-1</sup> CuCl<sub>2</sub>·2 H<sub>2</sub>O)<sup>155</sup> and 16 mg thiamine. Subsequently, fed batch fermentation was started with a constant feed of 15 g h<sup>-1</sup> using 50% glycerol and 1% MgSO<sub>4</sub>. After 1 h, protein expression was induced by the addition of 0.2 mM IPTG. After 2 h, (*R*)-limonene was added to start the biotransformation. Due to the reported antimicrobial effect of (*R*)-limonene,<sup>156</sup> the substrate was added repeatedly in aliquots to a final concentration of 5 mM. At different time points, samples were taken from the biotransformation to monitor substrate depletion and product formation by GC-FID analysis and measure acetate levels by HPLC analysis.

To determine the isolated product yields, the fermentation broth was extracted four times with ethyl acetate after addition of NaCl and the combined organic fractions were concentrated *in vacuo* and subjected to flash column chromatography.

### 2.3.8. Microbial toxicity of (1*R*,5*S*)-carveol on *E. coli* JM109

To determine a toxic effect of the biotransformation product (1*R*,5*S*)-carveol on growing *E. coli* JM109 cells, cultures were exposed to varying amounts of (1*R*,5*S*)-carveol (0 mM, 1 mM, 5 mM and 10 mM). *E. coli* JM109 cells carrying the plasmid encoding CDO M232A were inoculated in TB medium supplemented with 50 µg mL<sup>-1</sup> ampicillin to a start OD<sub>600</sub> of 0.05. After addition of carveol to the medium, cultures were grown at 30°C and 180 rpm. Cell growth was monitored by measuring the OD<sub>600</sub> and compared to growth of *E. coli* JM109 (pIP107D M232A) without carveol.

### 2.3.9. GC, HPLC and NMR analytics

#### Nuclear magnetic resonance (NMR) spectroscopy

<sup>1</sup>H and <sup>13</sup>C nuclear magnetic resonance (NMR) spectra were recorded on a Bruker Avance 500 spectrometer operating at 500.15 MHz and 125.76 MHz, respectively. All spectra were recorded at room temperature in CDCl<sub>3</sub>. Chemical shifts (δ) are expressed in parts per million (ppm) relative to tetramethylsilane (TMS, δ = 0 ppm). For the <sup>1</sup>H NMR spectra, data are reported as follows: chemical shift, integration, multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, br s = broad singlet, br t: broad triplet) and coupling constant (Hz). Assignments of chemical shifts were performed by heteronuclear single quantum coherence (HSQC) and heteronuclear multiple bond correlation (HMBC) analyses.

#### Gas chromatography (GC) analytics

GC-MS analyses were performed on a Shimadzu GCMS-QP2010 system equipped with an AOC-5000 auto injector. For product identification, GC-MS analyses were performed on a ZB-5 capillary column (Phenomenex, 15 m x 0.25 mm x 0.25 µm) or a DB-5 capillary column (Hewlett Packard, 30 m x 0.25 mm x 0.25 µm) using He as carrier gas (linear velocity 30 cm s<sup>-1</sup>) and split injection with an injector temperature of 250°C. Mass spectra were determined by EI ionization with 70 eV, an ion source temperature of 200°C and an interface temperature of 250°C. Mass detection was performed in scan mode from 40 *m/z* to 450 *m/z*.

The temperature program was as following:

2 min at 50°C, 10°C min<sup>-1</sup> to 280°C, hold for 2 min, 40°C min<sup>-1</sup> to 310°C and hold for 2 min

GC-FID analyses were performed on a Shimadzu GC-2010 system equipped with an AOC-20i auto injector. For quantification of product formation by GC-FID, a DB-5 capillary column (Hewlett Packard, 30 m x 0.25 mm x 0.25 µm) was used with H<sub>2</sub> as carrier gas (linear velocity 30 cm s<sup>-1</sup>). The injector temperature was 250°C and compounds were detected *via* flame ionization detector (FID) at 330°C. For monitoring of the product formation and substrate depletion during the up-scaling of (*R*)-limonene hydroxylation in a bioreactor, a Zebtron™ ZB-FFAP capillary column (Phenomenex, 30 m x 0.25 mm x 0.25 µm) was used with H<sub>2</sub> as carrier gas (linear velocity 30 cm s<sup>-1</sup>). The injector temperature was 250°C and compounds were detected *via* FID at 270°C.

The temperature programs for the analyses of the different compounds are described in Table 8.

**Table 8:** GC-FID temperature programs for analysis of biotransformations with substrates **1-(R)-7** performed on a DB-5 or a ZB-FFAP capillary column.

Compound	Column	Temperature program
styrene <b>1</b>	DB-5	3 min at 120°C, 10°C min <sup>-1</sup> to 140°C, 10°C min <sup>-1</sup> to 178°C, 75°C min <sup>-1</sup> to 310°C and hold for 4 min
vinylcyclohexane <b>2</b>	DB-5	3 min at 120°C, 10°C min <sup>-1</sup> to 140°C, 10°C min <sup>-1</sup> to 178°C, 75°C min <sup>-1</sup> to 310°C and hold for 4 min
cyclohexene <b>3</b>	DB-5	3 min at 70°C, 4°C min <sup>-1</sup> to 85°C, 10°C min <sup>-1</sup> to 190°C, 75°C min <sup>-1</sup> to 310°C and hold for 4 min
indene <b>4</b>	DB-5	3 min at 120°C, 10°C min <sup>-1</sup> to 140°C, 10°C min <sup>-1</sup> to 178°C, 75°C min <sup>-1</sup> to 310°C and hold for 4 min
myrcene <b>5</b>	DB-5	3 min at 120°C, 10°C min <sup>-1</sup> to 145°C, 5°C min <sup>-1</sup> to 178°C, 75°C min <sup>-1</sup> to 310°C and hold for 4 min
(+)- $\alpha$ -pinene <b>6</b>	DB-5	3 min at 120°C, 10°C min <sup>-1</sup> to 140°C, 10°C min <sup>-1</sup> to 178°C, 75°C min <sup>-1</sup> to 310°C and hold for 4 min
(R)-limonene (R)- <b>7</b>	DB-5	3 min at 120°C, 10°C min <sup>-1</sup> to 140°C, 10°C min <sup>-1</sup> to 178°C, 75°C min <sup>-1</sup> to 310°C and hold for 4 min
	ZB-FFAP <sup>[a]</sup>	3 min at 70°C, 10°C min <sup>-1</sup> to 100°C, 10°C min <sup>-1</sup> to 150°C, 40°C min <sup>-1</sup> to 230°C and hold for 4 min

<sup>[a]</sup> Analysis of (R)-limonene (R)-**7** and its biotransformation products was performed without derivatization on the ZB-FFAP capillary column.

For quantification of product formation by GC-FID, standard curves of the respective products were used with 1-decanol as internal standard. Standard curves were prepared in Eppendorf tubes as dilutions in 100 mM KPi pH 7.2. The respective products were prepared as 1 M dimethyl sulfoxide (DMSO) stocks and diluted to 20 mM in 100 mM KPi for the preparation of a concentration range from 0.1 mM–10 mM or 12 mM. Extraction with ethyl acetate and further work up was performed as described for whole cell biotransformations.

For GC-FID analyses on the DB-5 capillary column, sample derivatization was performed with an excess of acetic anhydride and pyridine (50  $\mu$ l sample was mixed with 40  $\mu$ l of each pyridine and acetic anhydride), except for biotransformation products of (+)- $\alpha$ -pinene **6** that were quantified without derivatization. To ensure complete derivatization, samples were incubated for 1 h at 70°C except for biotransformation reactions of styrene **1**. Due to decomposition of the underivatized arene-1,2-dihydrodiol **1b** at high temperatures, samples were incubated for 2 h at room temperature.

Due to polymerization of 3-vinylcyclohexa-3,5-diene-1,2-diol **1b** when concentrated *in vacuo*, synthesis and purification of the arene-1,2-dihydrodiol to use it as standard could not be performed. Furthermore, during synthesis of 1-indenol **4a**, explosive hydroperoxide intermediates (indene-1-peroxide) are formed<sup>157</sup> and hence, relative response factor (RF) values were used to determine concentrations for cyclohexa-3,5-diene-1,2-diol **1b** and 1-indenol **4a** in reference to a standard (1-phenyl-1,2-ethanediol **1a** and 1-indanol, respectively). RF values were calculated from the effective carbon numbers (ECN) of the respective compounds (ECN<sub>x</sub>) in relation to the ECN of the standard (ECN<sub>STD</sub>) using equation (1).<sup>158,159</sup>

$$RF = \frac{ECN_x}{ECN_{STD}} \quad (1)$$

The concentrations of the analytes were determined by external calibration using standards and normalized with the calculated RF values (Table 9). Due to identical ECN values for 3,10-dihydro myrcene **5b**, 2-methyl-6-methyleneoct-7-ene-2,3-diol **5c** and 1,2-dihydro myrcene **5d** after derivatization, quantification of 2-methyl-6-methyleneoct-7-ene-2,3-diol **5c** and 1,2-dihydro myrcene **5d** was performed using 3,10-dihydro myrcene **5b** as standard. Likewise, the acetylated forms of (*Z*)-2-methyl-6-methyleneocta-2,7-dien-1-ol **5a** and 2-methyl-6-methylene-octa-1,7-dien-3-ol resulting from dehydration of 2-methyl-6-methyleneoct-7-ene-2,3-diol **5c**, had identical ECN values and thus, quantification of 2-methyl-6-methylene-octa-1,7-dien-3-ol and (*Z*)-2-methyl-6-methyleneocta-2,7-dien-1-ol **5a** was performed using (*E*)-2-methyl-6-methyleneocta-2,7-dien-1-ol as standard.

**Table 9:** Calculated ECN and RF values for quantification of product formation in GC-FID analysis.

Compound	ECN	RF
<b>1a</b> <sup>[a]</sup>	9.5	
<b>1b</b> <sup>[a]</sup>	9.2	0.968
<b>4a</b> <sup>[a]</sup>	9.65	
1-indano <sup>[a]</sup>	9.75	0.990
<b>5a</b> <sup>[a]</sup>	10.45	
2-methyl-6-methylene-octa-1,7-dien-3-ol <sup>[a]</sup>	10.45	1
<b>5b</b> <sup>[a]</sup>	11.3	
<b>5c</b> <sup>[a]</sup>	11.3	1
<b>5d</b> <sup>[a]</sup>	11.3	
<b>5c</b> <sup>[a]</sup>	11.3	1
<b>7a</b> <sup>[a]</sup>	10.55	
<b>7b</b> <sup>[a]</sup>	10.55	1

<sup>[a]</sup> ECN and RF values were calculated for the acetylated compounds.

Dependent on the compound, chiral GC-FID analysis was performed either on a Supelco  $\beta$ -Dex 225 column (Sigma-Aldrich, 30 m  $\times$  0.25 mm  $\times$  0.25  $\mu$ m), on a Supelco  $\beta$ -Dex 110 column (Sigma-Aldrich, 30 m  $\times$  0.25 mm  $\times$  0.25  $\mu$ m), on a Hydrodex- $\beta$ -6TBDM column (Macherey-Nagel, 25 m  $\times$  0.25 mm) or on a CP-Chirasil-Dex CB column (Agilent, 25 m  $\times$  0.25 mm  $\times$  0.25  $\mu$ m) with H<sub>2</sub> as carrier gas (linear velocity 30 cm s<sup>-1</sup>). The injector temperature was 230°C and compounds were detected via flame ionization detector (FID) at 250°C. The temperature programs were as described in Table 10. Chiral analyses were performed with underivatized samples.

**Table 10:** Temperature programs and columns used for chiral GC-FID analysis of products **1a-4a**, **4b**, **5b-5c**, **6a** and **7b**.

Compound	Column	Temperature program
1-phenyl-1,2-ethanediol <b>1a</b>	Hydrodex- $\beta$ -6TBDM	2 min at 90°C, 2°C min <sup>-1</sup> to 120°C, 1°C min <sup>-1</sup> to 150°C, 40°C min <sup>-1</sup> to 230°C and hold for 3 min
1-cyclohexylethane-1,2-diol <b>2a</b>	CP-Chirasil-Dex CB	1 min at 40°C, 10°C min <sup>-1</sup> to 130°C, hold for 1 min, 0.5°C min <sup>-1</sup> to 190°C and hold for 1 min
2-cyclohexen-1-ol <b>3a</b>	Supelco $\beta$ -Dex 225	2 min at 70°C, 2°C min <sup>-1</sup> to 95°C, 1°C min <sup>-1</sup> to 115°C, 5°C min <sup>-1</sup> to 150°C, 40°C min <sup>-1</sup> to 210°C and hold for 2 min
1-indenol <b>4a</b>	Supelco $\beta$ -Dex 225	2 min at 90°C, 2°C min <sup>-1</sup> to 120°C, 1°C min <sup>-1</sup> to 150°C, 40°C min <sup>-1</sup> to 210°C and hold for 2 min
3,10-dihydro myrcene <b>5b</b> 1,2-dihydro myrcene <b>5d</b>	Supelco $\beta$ -Dex 110	1 min at 40°C, 10°C min <sup>-1</sup> to 115°C, hold for 60 min, 0.5°C min <sup>-1</sup> to 130°C
2-methyl-6-methyleneoct-7-ene-2,3-diol <b>5c</b>	Supelco $\beta$ -Dex 110	1 min at 40°C, 2.5°C min <sup>-1</sup> to 190°C and hold for 1 min
pinanediol <b>6a</b>	CP-Chirasil-Dex CB	1 min at 40°C, 2.5°C min <sup>-1</sup> to 190°C and hold for 1 min
carveol <b>7b</b>	CP-Chirasil-Dex CB	1 min at 40°C, 2.5°C min <sup>-1</sup> to 190°C and hold for 1 min

### High-performance liquid chromatography (HPLC) analytics

HPLC analysis was performed on an Agilent 1200 series system equipped with a G1315D diode array detector (DAD), a G1362A refractive index detector (RID), a G1322A degasser, a G1311A quaternary pump, a G1316A temperature controlled column compartment and a G1329A autosampler unit.

Chiral stationary-phase HPLC analysis of 1-phenyl-1,2-ethanediol **1a** and 1-indenol **4a** was performed with a Chiralpak IB column (Daicel Chiral Technologies Europe S.A.S., 5  $\mu$ m, 150 x 4.6 mm). For separation of 1-phenyl-1,2-ethanediol **1a** enantiomers, a mobile phase of hexane:isopropanol (93:7) at a flow rate of 0.9 mL min<sup>-1</sup> and 20°C was used. After injecting 10-20  $\mu$ l sample, the UV signal was monitored at the maximum specific absorbance wavelength (258 nm). Enantiomers of 1-indenol **4a** were separated with a mobile phase of hexane:isopropanol (90:10) at a flow rate of 1 mL min<sup>-1</sup> and 25°C. After injecting 10-20  $\mu$ l sample, the UV signal was monitored at the maximum specific absorbance wavelength (266 nm).

For chiral stationary-phase HPLC analysis of *cis*-1,2-indandiol **4b**, a Reprosil-AM column (Dr. Maisch, 5  $\mu$ m, 4.6 mm x 250 mm) was used. Separation of enantiomers was performed with a mobile phase of hexane:isopropanol (92:8) at a flow rate of 1 mL min<sup>-1</sup> and 25°C. 10  $\mu$ l sample were injected and the UV signal was monitored at 220 nm.

All chiral analyses were performed with underivatized samples.

For quantification of acetate levels during fed-batch fermentation, samples were centrifuged and the supernatant was filtered (0.45  $\mu$ m). Samples were measured on an Aminex HPX-87H Ion Exclusion Column (BIO-RAD, USA, 300 mm x 78 mm) at 60°C with a mobile phase of 5 mM sulfuric acid and a flow rate of 0.5 mL min<sup>-1</sup>. 10  $\mu$ l sample were injected and signals were monitored on a RID at 35°C and a DAD at 210 nm. Acetate concentrations were determined by external calibration using acetate standard curves (0.25 mM-40 mM) that were prepared in the same manner as the samples.

## 2.3.10. Chemical synthesis of product standards

In chemical synthesis, all non-aqueous reactions were run under an inert nitrogen atmosphere by using standard techniques for manipulating air-sensitive compounds and all glassware was flame dried prior to use.

Flash column chromatography was performed with 0.040-0.063 mm mesh silica. Analytical thin-layer chromatography (TLC) was performed on silica plates (Macherey-Nagel) and TLC analysis was visualized using anisaldehyde staining solution (50 mL glacial acetic acid, 0.5 mL *p*-anisaldehyde and 1 mL sulfuric acid 95-97%).

Synthesis of (*S*)- and (*R*)-1-cyclohexylethane-1,2-diol (**2a**)

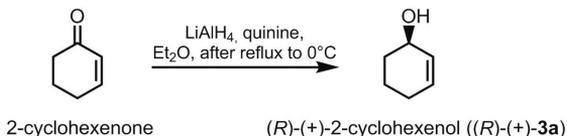
1.4 g AD-mix  $\alpha$  and  $\beta$ , respectively, was dissolved in a mix of *tert*-butanol and water (1:1, 5 mL each) at room temperature. After cooling the solution to 4°C, 1 mmol vinylcyclohexane was added and the suspension was stirred over night. 0.5 g Sodium metabisulfite was added slowly to the reaction mixture, followed by 20 mL of water and the suspension was stirred for 30 min at room temperature. Extraction was performed 3 x with ethyl acetate and the organic phases were combined, dried over anhydrous sodium sulfate and concentrated *in vacuo*. The product was purified by flash column chromatography (step gradient cyclohexane/EtOAc 1:2 and cyclohexane/EtOAc 1:3).

(*S*)-1-cyclohexylethane-1,2-diol ((*S*)-**2a**, AD-mix  $\alpha$ )

Colorless solid (90% yield); <sup>1</sup>H NMR  $\delta$  (CDCl<sub>3</sub>, 500 MHz), 0.98-1.09 (2 H, m), 1.10-1.29 (3 H, m), 1.36-1.44 (1 H, m), 1.60-1.70 (2 H, m), 1.70-1.80 (2 H, m), 1.87 (1 H, d, *J* = 12.51 Hz), 2.71 (2 H, br s), 3.40-3.46 (1 H, m), 3.49-3.55 (1 H, m), 3.69 (1 H, dd, *J*<sub>1</sub> = 11.05 Hz, *J*<sub>2</sub> = 2.83 Hz); <sup>13</sup>C NMR  $\delta$  (CDCl<sub>3</sub>, 500 MHz), 25.99, 26.07, 26.38, 28.66, 28.95, 40.70, 64.80, 76.52. MS (GC, EI): *m/z* (%) = 127 (0.05, M<sup>+</sup> - H<sub>2</sub>O), 113 (34), 95 (100), 67 (21), 55 (19), 41 (17)

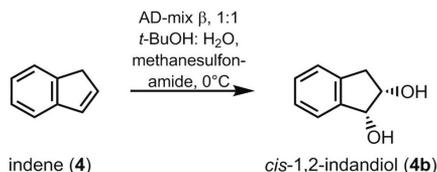
(*R*)-1-cyclohexylethane-1,2-diol ((*R*)-**2a**, AD-mix  $\beta$ )

Colorless solid (86% yield); <sup>1</sup>H NMR  $\delta$  (CDCl<sub>3</sub>, 500 MHz), 0.98-1.09 (2 H, m), 1.10-1.30 (3 H, m), 1.35-1.45 (1 H, m), 1.60-1.70 (2 H, m), 1.70-1.80 (2 H, m), 1.87 (1 H, d, *J* = 12.53 Hz), 2.61 (2 H, br s), 3.40-3.47 (1 H, m), 3.48-3.55 (1 H, m), 3.70 (1 H, d, *J* = 10.64 Hz); <sup>13</sup>C NMR  $\delta$  (CDCl<sub>3</sub>, 500 MHz), 25.99, 26.07, 26.37, 28.66, 28.95, 40.70, 64.80, 76.51. MS (GC, EI): *m/z* (%) = 127 (0.06, M<sup>+</sup> - H<sub>2</sub>O), 113 (33), 95 (100), 67 (24), 55 (20), 41 (20)

Synthesis of (*R*)-(+)-2-cyclohexenol ((*R*)-(+)-**3a**)

Synthesis of (*R*)-(+)-2-cyclohexenol was performed according to Wynberg *et al.*<sup>160</sup> yielding an enantiomeric excess of 13% as determined by chiral GC-FID analysis. To 0.56 g (15 mmol) LiAlH<sub>4</sub> in 100 mL dry diethyl ether, 4.75 g (15 mmol) quinine was added and the suspension was refluxed for 15 min. After cooling to 0°C, 1.28 g (13 mmol) cyclohexenone dissolved in 3.3 mL diethyl ether were added and the mixture was stirred for 1 h at 0°C. After addition of 2 mL water and 10% H<sub>2</sub>SO<sub>4</sub>, the reaction was extracted 3 x with 50 mL ethyl acetate and the combined organic phases were dried over anhydrous sodium sulfate and concentrated *in vacuo*. The product was purified by flash column chromatography (cyclohexane/EtOAc 3:1).

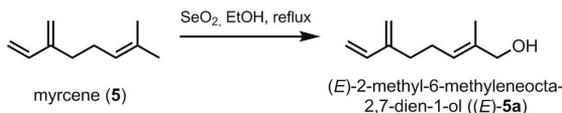
Colorless oil (2% yield); <sup>1</sup>H NMR δ (CDCl<sub>3</sub>, 500 MHz), 1.51-1.66 (3 H, m), 1.68-1.79 (1 H, m), 1.83-1.91 (1 H, m), 1.92-2.01 (2 H, m), 4.20 (1 H, s), 5.73-5.78 (1 H, m), 5.81-5.86 (1 H, m); <sup>13</sup>C NMR δ (CDCl<sub>3</sub>, 500 MHz), 18.93, 25.06, 32.01, 65.51, 129.85, 130.60. MS (GC, EI): *m/z* (%) = 99 (2, M<sup>+</sup>), 98 (28, M), 97 (28), 83 (50), 79 (28), 77 (16), 70 (100), 69 (46), 55 (46), 43 (24)

Synthesis of (1*R*,2*S*)-*cis*-1,2-indandiol (**4b**)

1.4 g AD-mix β was dissolved in a mix of *tert*-butanol and water (1:1, 5 mL each) at room temperature and 95 mg methanesulfonamide was added to accelerate the reaction. After cooling the solution to 4°C, 1 mmol indene was added and the reaction mixture stirred over night. 0.5 g Sodium metabisulfite and 20 mL water were added. After warming to room temperature, the reaction mixture was stirred for 30 min and extracted three times with ethyl acetate. The organic phases were combined, dried over anhydrous sodium sulfate and concentrated *in vacuo*. The product was purified by flash column chromatography (cyclohexane/EtOAc 1:1).

White solid (69% yield);  $^1\text{H NMR } \delta$  ( $\text{CDCl}_3$ , 500 MHz), 2.90 (1 H, dd,  $J_1 = 16.31$  Hz,  $J_2 = 3.49$  Hz), 2.94-3.01 (1 H, m), 3.06 (1 H, dd,  $J_1 = 16.31$  Hz,  $J_2 = 5.66$  Hz), 3.09-3.15 (1 H, m), 4.34-4.44 (1 H, m), 4.87-4.95 (1H, m), 7.19-7.28 (3H, m), 7.37-7.41 (1H, m);  $^{13}\text{C NMR } \delta$  ( $\text{CDCl}_3$ , 500 MHz), 38.51, 73.43, 75.95, 125.05, 125.33, 127.14, 128.77, 140.15, 141.94. MS (GC, EI):  $m/z$  (%) = 151 (3,  $\text{M}^+$ ), 150 (29, M), 119 (13), 107 (64), 104 (100), 103 (44), 91 (34), 79 (26), 77 (41), 65 (15), 63 (9), 51 (18)

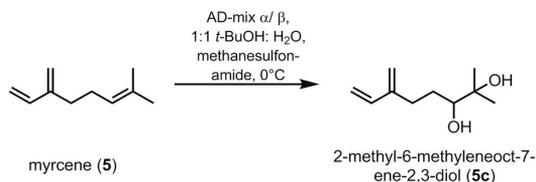
### Synthesis of (*E*)-2-methyl-6-methyleneocta-2,7-dien-1-ol ((*E*)-**5a**)



Synthesis of (*E*)-2-methyl-6-methyleneocta-2,7-dien-1-ol was performed according to Granger *et al.*<sup>161</sup> To a solution of 7 g (0.05 mol) myrcene in 50 mL pure EtOH, 3 g (0.025 mol)  $\text{SeO}_2$  were added and the reaction mixture was refluxed for 1 h. After 3 x extraction with 50 mL diethyl ether, organic phases were combined, dried over anhydrous sodium sulfate and concentrated *in vacuo*. The product was purified by flash column chromatography (cyclohexane/EtOAc 5:1).

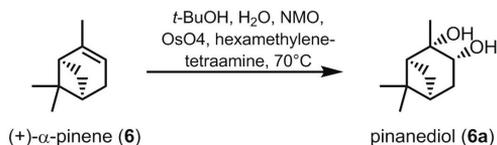
Yellow oil (14% yield);  $^1\text{H NMR } \delta$  ( $\text{CDCl}_3$ , 500 MHz), 1.67 (3 H, s), 2.26 (4 H, br s), 4.00 (2 H, s), 5.01 (2 H, d,  $J = 14.10$  Hz), 5.07 (1 H, d,  $J = 10.59$  Hz), 5.24 (1 H, d,  $J = 17.61$  Hz), 5.41-5.47 (1 H, m), 6.38 (1 H, dd,  $J_1 = 17.58$  Hz,  $J_2 = 10.86$  Hz);  $^{13}\text{C NMR } \delta$  ( $\text{CDCl}_3$ , 500 MHz), 13.70, 26.20, 31.02, 68.86, 113.18, 115.88, 125.66, 135.15, 138.87, 145.81. MS (GC, EI):  $m/z$  (%) = 152 (0.16, M), 134 (20), 121 (12), 119 (29), 105 (16), 93 (94), 91 (40), 80 (10), 79 (48), 77 (24), 67 (25), 55 (18), 43 (100)

### Synthesis of 2-methyl-6-methyleneoct-7-ene-2,3-diol (**5c**)



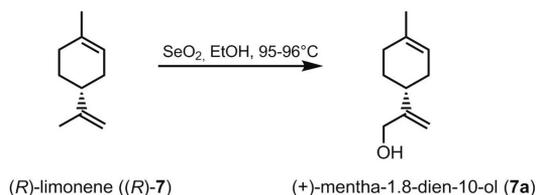
1.4 g AD-mix  $\alpha$  or  $\beta$  was dissolved in a mix of *tert*-butanol and water (1:1, 5 mL each) at room temperature and 95 mg methanesulfonamide was added to accelerate the reaction. After cooling the solution to 4°C, 1 mmol myrcene was added and the reaction mixture stirred over night. 0.5 g Sodium metabisulfite and 20 mL water were added. After warming to room temperature, the reaction mixture was stirred for 30 min and extracted three times with ethyl acetate. The organic phases were combined, dried over anhydrous sodium sulfate and concentrated *in vacuo*. The product was purified by flash column chromatography (step gradient cyclohexane/EtOAc 1:1 and cyclohexane/EtOAc 1:2).



Synthesis of (1*S*,2*S*,3*R*,5*S*)-(+)-pinanediol (**6a**) with OsO<sub>4</sub>

Synthesis of (1*S*,2*S*,3*R*,5*S*)-(+)-pinanediol with OsO<sub>4</sub> was performed according to Erdik *et al.*<sup>163</sup> To a solution of 4.9 g (0.041 mol) NMO and 15.5 g (0.11 mol) hexamethylenetetraamine in 85 mL *tert*-butanol and 15 mL H<sub>2</sub>O, 5.6 g (0.04 mol) (+)- $\alpha$ -pinene and 0.14 g (0.55 mmol) OsO<sub>4</sub> were added under nitrogen atmosphere. The reaction mixture was heated to 70°C for 10 h and after cooling, 100 mL of a 10% aqueous sodium bisulfite solution was added. The reaction mixture was stirred until the dark brown color of the aqueous phase changed to yellow and the aqueous phase was extracted 3 x with 100 mL diethyl ether. The organic phases were combined, dried over anhydrous sodium sulfate and concentrated *in vacuo*. The product was purified by flash column chromatography (cyclohexane/EtOAc 1:2).

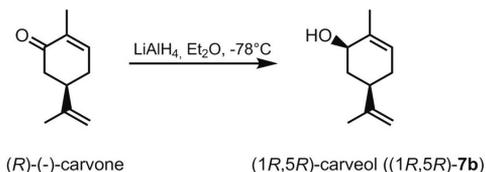
Colorless oil that started to solidify after some time (92% yield); <sup>1</sup>H NMR  $\delta$  (CDCl<sub>3</sub>, 500 MHz), 0.93 (3 H, s), 1.27 (3 H, s), 1.31 (3 H, s), 1.37 (1 H, d,  $J$  = 10.39 Hz), 1.61-1.67 (1 H, m), 1.89-1.94 (1 H, m), 2.01 (1 H, t,  $J$  = 5.8 Hz), 2.17-2.23 (1 H, m), 2.41-2.49 (1 H, m), 2.80 (1 H, br s), 3.12 (1 H, d,  $J$  = 6.58 Hz), 3.99 (1 H, ddd,  $J_1$  = 9.36 Hz,  $J_2$  = 6.58 Hz,  $J_3$  = 5.23 Hz); <sup>13</sup>C NMR  $\delta$  (CDCl<sub>3</sub>, 500 MHz), 24.13, 27.82, 28.04, 29.59, 38.11, 38.94, 40.50, 53.95, 69.16, 73.81. MS (GC, EI):  $m/z$  (%) = 137 (2), 126 (16), 111 (27), 99 (46), 93 (16), 83 (17), 81 (20), 71 (36), 55 (35), 43 (100)

Synthesis of (+)-mentha-1.8-dien-10-ol (**7a**)

Synthesis of (+)-mentha-1.8-dien-10-ol was performed according to Thomas and Bucher.<sup>164</sup> 11 g of (*R*)-limonene were warmed to 50°C and a hot solution of SeO<sub>2</sub> (4.4 g dissolved in 7 mL of EtOH) were added dropwise within 30 min. The reaction mixture was stirred for 2 h at 95-96°C and after cooling, the precipitated selenium was removed by filtration, washed with diethyl ether and EtOH. The combined organic phases were dried over anhydrous sodium sulfate and concentrated *in vacuo*. The product was purified by flash column chromatography (1. column: DCM/cyclohexene/EtOAc 20:4:0.5, 2. column: cyclohexene/EtOAc 10:1).

Yellow oil (2% yield);  $^1\text{H NMR } \delta$  ( $\text{CDCl}_3$ , 500 MHz), 1.43-1.56 (2 H, m), 1.65 (3 H, s), 1.78-2.23 (6 H, m), 4.14 (2 H, d,  $J = 5.95$  Hz), 4.91 (1 H, s), 5.04-5.07 (1 H, m), 5.40 (1 H, br s);  $^{13}\text{C NMR } \delta$  ( $\text{CDCl}_3$ , 500 MHz), 23.45, 28.18, 30.53, 31.34, 36.89, 65.19, 107.83, 120.43, 133.83, 153.65. MS (GC, EI):  $m/z$  (%) = 153 (1,  $\text{M}^+$ ), 152 (14, M), 134 (56), 119 (93), 106 (99), 105 (42), 93 (71), 91 (89), 79 (93), 67 (100), 55 (56)

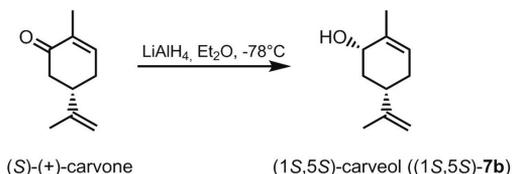
### Synthesis of (1*R*,5*R*)-carveol (**7b**)



Synthesis of (1*R*,5*R*)-carveol starting from (*R*)-(-)-carvone was performed according to Dhulut *et al.*<sup>165</sup> 2 g (13.3 mmol) (*R*)-(-)-carvone in 10 mL diethyl ether were slowly added to 6.7 mL (6.7 mmol) of a 1 M solution of  $\text{LiAlH}_4$  at  $-78^\circ\text{C}$ . After 30 min, 450  $\mu\text{l}$  (25 mmol)  $\text{H}_2\text{O}$ , 450  $\mu\text{l}$  (25 mmol)  $\text{NaOH}$  and additional 450  $\mu\text{l}$  (25 mmol)  $\text{H}_2\text{O}$  were successively added and the reaction mixture was stirred until a white precipitate appeared. The reaction mixture was dried over anhydrous sodium sulfate and concentrated *in vacuo*. The product was purified by flash column chromatography (cyclohexene/EtOAc 8:2).

Colorless oil (39% yield);  $^1\text{H NMR } \delta$  ( $\text{CDCl}_3$ , 500 MHz), 1.51 (2 H, td,  $J_1 = 12.10$  Hz,  $J_2 = 9.52$ ), 1.74 (3 H, br s), 1.76 (3 H, br s), 1.90-2.33 (4 H, m), 4.19 (1 H, br s), 4.73 (2 H, s), 5.50 (1 H, br s);  $^{13}\text{C NMR } \delta$  ( $\text{CDCl}_3$ , 500 MHz), 18.99, 20.66, 31.02, 37.99, 40.42, 70.91, 109.16, 123.90, 136.13, 149.00. MS (GC, EI):  $m/z$  (%) = 153 (0.2,  $\text{M}^+$ ), 152 (1, M), 134 (54), 119 (32), 109 (62), 91 (39), 84 (100), 69 (51), 55 (70)

### Synthesis of (1*S*,5*S*)-carveol (**7b**)



Synthesis of (1*S*,5*S*)-carveol starting from (*S*)-(+)-carvone was performed according to Bermejo *et al.*<sup>166</sup> 2 g (13.3 mmol) (*S*)-(+)-carvone in 10 mL diethyl ether were slowly added to 6.7 mL

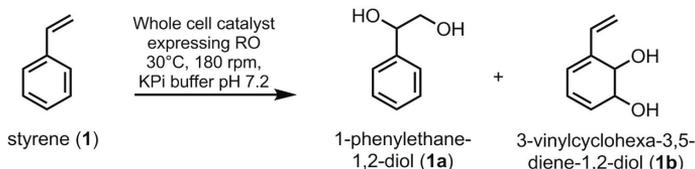
(6.7 mmol) of a 1 M solution of  $\text{LiAlH}_4$  at  $-78^\circ\text{C}$ . After 1.5 h, 450  $\mu\text{l}$  (25 mmol)  $\text{H}_2\text{O}$ , 450  $\mu\text{l}$  (25 mmol)  $\text{NaOH}$  and additional 450  $\mu\text{l}$  (25 mmol)  $\text{H}_2\text{O}$  were successively added and the reaction mixture was stirred until a white precipitate appeared. The reaction mixture was dried over anhydrous sodium sulfate and concentrated *in vacuo*. The product was purified by flash column chromatography (cyclohexene/EtOAc 8:2).

Colorless oil (40% yield);  $^1\text{H NMR}$   $\delta$  ( $\text{CDCl}_3$ , 500 MHz), 1.51 (1 H, td,  $J_1 = 12.22$ ,  $J_2 = 9.53$ ), 1.60 (1 H, br s), 1.74 (3 H, s), 1.76 (3 H, br s), 1.90-2.32 (4 H, m), 4.19 (1 H, br s), 4.73 (2 H, s), 5.50 (1 H, br s);  $^{13}\text{C NMR}$   $\delta$  ( $\text{CDCl}_3$ , 500 MHz), 19.00, 20.66, 31.03, 37.99, 40.44, 70.91, 109.16, 123.88, 136.15, 149.00. MS (GC, EI):  $m/z$  (%) = 153 (0.1,  $\text{M}^+$ ), 152 (1, M), 134 (39), 119 (29), 109 (54), 91 (48), 84 (82), 69 (47), 55 (73)

### 2.3.11. Identification of biotransformation products by NMR and GC-MS analysis

Preparative large scale biotransformations for product isolation were performed only with the RO enzyme showing the best conversion rates. In case the amount of formed product was too low to be isolated, product identification was performed using an authentic standard in combination with GC-MS analysis.

Biotransformation of styrene (**1**) to 1-phenyl-1,2-ethanediol (**1a**) and 3-vinylcyclohexa-3,5-diene-1,2-diol (**1b**)



#### 1-phenyl-1,2-ethanediol (**1a**)

RO-catalyzed biotransformations

MS (GC, EI)  $m/z$  (%) = 139 (0.9,  $\text{M}^+$ ), 138 (10, M), 107 (100), 79 (66), 77 (40), 51 (10)

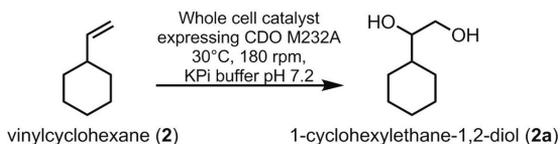
Standard Fluka

MS (GC, EI)  $m/z$  (%) = 139 (0.9,  $\text{M}^+$ ), 138 (10, M), 107 (100), 79 (82), 77 (51), 51 (12)

#### 3-vinylcyclohexa-3,5-diene-1,2-diol (**1b**)

MS (GC, EI)  $m/z$  (%) = 139 (1,  $\text{M}^+$ ), 138 (11, M), 120 (44), 109 (13), 92 (40), 91 (100), 81 (23), 79 (27), 77 (27), 65 (22), 53 (18)

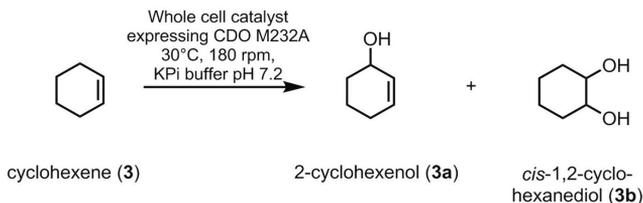
Biotransformation of vinylcyclohexane (**2**) to 1-cyclohexylethane-1,2-diol (**2a**) with CDO M232A



The product was purified by flash column chromatography (cyclohexane/EtOAc 1:3).

$^1\text{H NMR } \delta$  ( $\text{CDCl}_3$ , 500 MHz), 0.97-1.08 (2 H, m), 1.09-1.29 (3 H, m), 1.34-1.44 (1 H, m), 1.65 (2 H, t,  $J = 13.16$  Hz), 1.69-1.79 (2 H, m), 1.87 (1 H, d,  $J = 12.62$  Hz), 3.38-3.44 (1 H, m), 3.47-3.52 (1 H, m), 3.54 (1 H, br s), 3.67 (1 H, dd,  $J_1 = 11.20$  Hz,  $J_2 = 2.50$  Hz);  $^{13}\text{C NMR } \delta$  ( $\text{CDCl}_3$ , 500 MHz), 26.03, 26.11, 26.40, 28.70, 29.01, 40.71, 64.75, 76.53. MS (GC, EI):  $m/z$  (%) = 127 (0.05,  $\text{M}^+ - \text{H}_2\text{O}$ ), 113 (34), 95 (100), 67 (21), 55 (19), 41 (17)

Biotransformation of cyclohexene (**3**) to 2-cyclohexenol (**3a**) and *cis*-1,2-cyclohexanediol (**3b**)



2-cyclohexenol (**3a**)

RO-catalyzed biotransformations

MS (GC, EI)  $m/z$  (%) = 99 (2,  $\text{M}^+$ ), 98 (31, M), 97 (34), 83 (51), 79 (20), 70 (100), 69 (32), 55 (38), 41 (36)

Standard Fluka

MS (GC, EI)  $m/z$  (%) = 99 (2,  $\text{M}^+$ ), 98 (30, M), 97 (34), 83 (52), 79 (31), 70 (100), 69 (31), 55 (37), 41 (33)

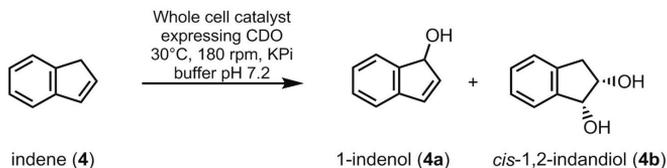
*cis*-1,2-cyclohexanediol (**3b**)

RO-catalyzed biotransformations

MS (GC, EI)  $m/z$  (%) = 116 (8, M), 98 (34), 83 (33), 70 (100), 69 (22), 57 (61), 42 (31), 41 (37)

Standard Sigma-Aldrich

MS (GC, EI)  $m/z$  (%) = 117 (0.6,  $\text{M}^+$ ), 116 (8, M), 98 (35), 83 (34), 70 (100), 69 (21), 57 (61), 42 (31), 41 (36)

Biotransformation of indene (**4**) to 1-indenol (**4a**) and *cis*-1,2-indandiol (**4b**) with CDO

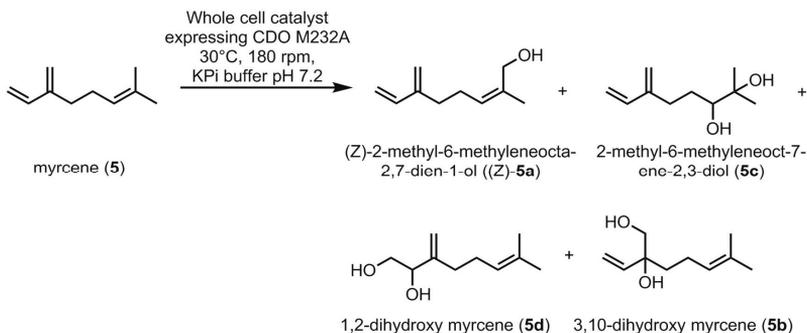
The product was purified by flash column chromatography (step gradient cyclohexane/DCM/EtOAc 4:4:0.5 and cyclohexane/EtOAc 1:2).

1-indenol (**4a**)

Slightly yellow oil;  $^1\text{H NMR } \delta$  ( $\text{CDCl}_3$ , 500 MHz), 1.71 (1 H, br s), 5.17 (1 H, s), 6.40 (1 H, dd,  $J_1 = 5.96$  Hz,  $J_2 = 1.88$  Hz), 6.73 (1 H, d,  $J = 5.63$  Hz), 7.18-7.29 (3 H, m), 7.51 (1 H, d,  $J = 7.15$  Hz);  $^{13}\text{C NMR } \delta$  ( $\text{CDCl}_3$ , 500 MHz), 77.59, 121.42, 123.42, 126.13, 128.51, 132.72, 137.70, 142.32, 145.45. MS (GC, EI):  $m/z$  (%) = 133 (9,  $\text{M}^+$ ), 132 (100, M), 131 (87), 115 (15), 103 (42), 89 (5), 77 (28), 63 (8), 51 (18)

*cis*-1,2-indandiol (**4b**)

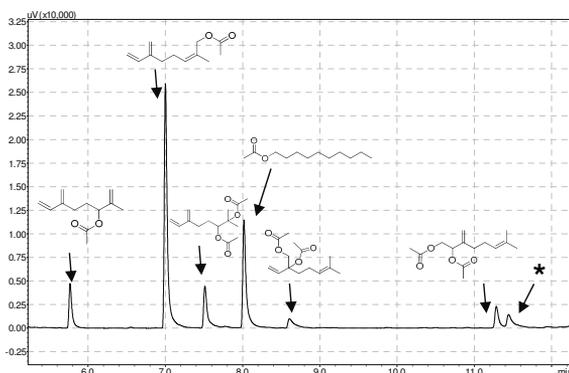
White solid;  $^1\text{H NMR } \delta$  ( $\text{CDCl}_3$ , 500 MHz), 2.86 (1 H, dd,  $J_1 = 16.31$  Hz,  $J_2 = 3.45$  Hz), 3.01 (1 H, dd,  $J_1 = 16.31$  Hz,  $J_2 = 5.74$  Hz), 3.30 (1 H, br s), 3.45 (1 H, br s), 4.27-4.37 (1 H, m), 4.80-4.89 (1 H, m), 7.17-7.27 (3 H, m), 7.34-7.38 (1 H, m);  $^{13}\text{C NMR } \delta$  ( $\text{CDCl}_3$ , 500 MHz), 38.40, 73.41, 75.93, 125.05, 125.28, 127.08, 128.71, 140.16, 141.94. MS (GC, EI):  $m/z$  (%) = 151 (3,  $\text{M}^+$ ), 150 (28, M), 119 (14), 107 (62), 104 (100), 103 (44), 91 (35), 79 (27), 77 (42), 65 (16), 63 (9), 51 (19)

Biotransformation of myrcene (**5**) to (*Z*)-2-methyl-6-methyleneocta-2,7-dien-1-ol ((*Z*)-**5a**), 3,10-dihydro myrcene (**5b**), 2-methyl-6-methyleneoct-7-ene-2,3-diol (**5c**) and 1,2-dihydro myrcene (**5d**) with CDO M232A



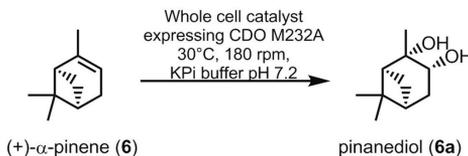
Slightly yellow oil;  $^1\text{H NMR } \delta$  ( $\text{CDCl}_3$ , 500 MHz), 1.62 (1 H, br s), 1.74 (3 H, s), 1.69-1.81 (2 H, m), 2.19-2.27 (1 H, m), 2.28-2.38 (1 H, m), 4.11 (1 H, t,  $J = 6.43$  Hz), 4.85-4.87 (1 H, m), 4.97 (1 H, s), 5.03 (2 H, br s), 5.06 (1 H, d,  $J = 10.91$  Hz), 5.25 (1 H, d,  $J = 17.72$  Hz), 6.38 (1 H, dd,  $J_1 = 10.63$  Hz,  $J_2 = 17.63$  Hz);  $^{13}\text{C NMR } \delta$  ( $\text{CDCl}_3$ , 500 MHz), 17.65, 27.32, 33.37, 75.62, 111.14, 113.40, 115.87, 138.76, 146.02, 147.44. MS (GC, EI):  $m/z$  (%) = 152 (0.39, M), 137 (13), 134 (8), 123 (20), 119 (23), 109 (24), 105 (23), 96 (19), 93 (27), 91 (30), 84 (37), 81 (23), 79 (49), 71 (58), 69 (72), 67 (77), 55 (32), 53 (35), 43 (93), 41 (100)

During the CDO M232A-catalyzed biotransformation of **5**, a potential byproduct (\*) might be formed in addition to the above described products which could not be identified due to low amounts formed (estimated < 5%).



**Figure 7:** GC-FID chromatogram of the CDO M232A-catalyzed biotransformation of myrcene **5** after derivatization. (\*): potential byproduct.

### Biotransformation of (+)- $\alpha$ -pinene (**6**) to pinanediol (**6a**) with CDO M232A



The product was purified by flash column chromatography (step gradient cyclohexane/EtOAc 1:1 and cyclohexane/EtOAc 1:2).

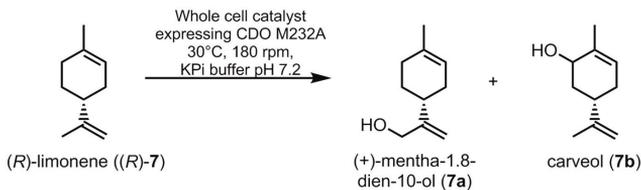
$^1\text{H NMR } \delta$  ( $\text{CDCl}_3$ , 500 MHz), 0.94 (3 H, s), 1.27 (3 H, s), 1.31 (3 H, s), 1.37 (1 H, d,  $J = 10.39$  Hz), 1.61-1.67 (1 H, m), 1.89-1.94 (1 H, m), 2.01 (1 H, t,  $J = 5.8$  Hz), 2.17-2.23 (1 H, m), 2.42-2.49 (1 H, m), 2.77 (1 H, br s), 3.09 (1 H, br s), 3.99 (1 H, dd,  $J_1 = 9.36$  Hz,  $J_2 = 5.20$  Hz);  $^{13}\text{C NMR } \delta$  ( $\text{CDCl}_3$ , 500 MHz), 24.13, 27.82, 28.03, 29.56, 38.10, 38.95, 40.50, 53.96, 69.19, 73.86. MS (GC, EI):  $m/z$  (%) = 137 (2), 126 (15), 111 (26), 99 (42), 93 (16), 83 (17), 81 (20), 71 (34), 55 (35), 43 (100)

During the CDO M232A-catalyzed biotransformation of **6**, a potential byproduct (\*) might be formed in addition to the above described product which could not be identified due to low amounts formed (estimated < 5%).



**Figure 8:** GC-FID chromatogram of the CDO M232A-catalyzed biotransformation of (+)- $\alpha$ -pinene **6**. (\*): potential byproduct.

Biotransformation of (*R*)-limonene ((*R*)-**7**) to (+)-mentha-1,8-dien-10-ol (**7a**) and carveol (**7b**) with CDO M232A



(+)-mentha-1,8-dien-10-ol (**7a**)

MS (GC, EI):  $m/z$  (%) = 152 (9, M), 134 (49), 119 (89), 106 (91), 105 (38), 93 (66), 91 (89), 79 (87), 67 (100), 55 (69)

Carveol was purified by flash column chromatography (step gradient cyclohexane/EtOAc 8:1 and cyclohexane/EtOAc 6:1).



**isopiperitenol (7c)**

Colorless oil;  $^1\text{H NMR } \delta$  ( $\text{CDCl}_3$ , 500 MHz), 1.55-1.65 (1 H, m), 1.70 (3 H, s), 1.70-1.73 (1 H, m), 1.73 (3 H, s), 1.82 (1 H, s), 1.89-1.97 (1 H, m), 2.03-2.12 (2 H, m), 4.08-4.16 (1 H, m), 4.87 (2 H, d,  $J = 19.36$  Hz), 5.45 (1 H, s);  $^{13}\text{C NMR } \delta$  ( $\text{CDCl}_3$ , 500 MHz), 19.36, 23.07, 26.21, 30.23, 50.95, 68.71, 112.36, 124.34, 136.75, 146.47. MS (GC, EI):  $m/z$  (%) = 152 (2, M), 134 (2), 119 (3), 109 (8), 108 (9), 91 (6), 84 (100), 83 (46), 81 (10), 77 (6), 69 (14), 56 (20), 53 (8), 43 (5)

**1-methyl-4-(prop-1-en-2-yl) cyclohex-2-enol (7d)**

Standard (Sigma-Aldrich)

MS (GC, EI):  $m/z$  (%) = 137 (51), 134 (93), 119 (50), 109 (100), 95 (35), 93 (30), 91 (56), 81 (33), 79 (74), 77 (35), 69 (24), 67 (35), 55 (23), 43 (81), 41 (37)

CDO I288V-catalyzed biotransformation

MS (GC, EI):  $m/z$  (%) = 152 (3, M), 137 (56), 134 (94), 119 (44), 109 (100), 95 (36), 93 (29), 91 (44), 81 (34), 79 (76), 77 (33), 69 (26), 67 (36), 55 (24), 43 (89), 41 (38)

**2.3.12. Docking**

Docking of styrene **1** and indene **4** into the active sites of CDO wild type and variants was performed using AMBER force field-03<sup>167</sup> in implementation of AutoDock Vina in Yasara.<sup>168</sup>

### 3. RESULTS

#### 3.1. Oxyfunctionalization of alkenes by wild type ROs

ROs are naturally involved in the degradation of aromatic compounds and a vast array of arenes is known to be accepted by these enzymes.<sup>65</sup> Despite their broad substrate specificity, the conversion of unnatural substrates has only started to receive a growing interest within recent years and limited data is available concerning the RO-catalyzed oxidation of olefins when compared to arene dihydroxylation. To explore the potential of ROs for the oxyfunctionalization of alkenes, NDO from *Pseudomonas* sp. NCIB 9816-4, BDO from *P. putida* ML2 and CDO from *Pseudomonas fluorescens* IP01 were examined for their ability to oxidize aliphatic C=C double bonds. The three enzymes were selected as they are amongst the ROs that have already been reported to catalyze the oxyfunctionalization of several olefins.<sup>73,74,92,146,169</sup> Furthermore, the crystal structures of NDO (PDB code: 1O7N)<sup>101</sup> and CDO (PDB code: 1WQL)<sup>170</sup> have been solved, providing information of the enzymes' active sites.<sup>101,170</sup> For BDO, no structural data were available, yet, due to its high sequence identity of 93% with the  $\alpha$ -subunit of TDO from *P. putida* F1,<sup>153</sup> a homology model could be generated based on the TDO crystal structure (PDB code: 3EN1).<sup>171</sup>

As ROs are cofactor-dependent multicomponent enzymes,<sup>76</sup> *in vivo* biotransformations were performed using 0.2 g mL<sup>-1</sup> (*cww*) resting *E. coli* cells supplemented with glucose for *in situ* cofactor regeneration. In the whole cell system, oxygenase subunits as well as electron transport components (reductase and ferredoxin) were simultaneously expressed in one cell, also addressing the challenge of low stability reported for purified oxygenase enzymes.<sup>172,173</sup> As negative control, *E. coli* harboring the empty vector was employed to exclude product formation catalyzed by endogenous enzymes.

##### 3.1.1. Cloning and expression of RO genes

Plasmids encoding NDO, BDO and CDO genes for the simultaneous expression of all RO components in *E. coli* were obtained from Rebecca Parales, David Leak and Hideaki Nojiri. For NDO and BDO, expression was performed as reported in literature<sup>145,153</sup> and the presence of functional enzymes was confirmed with the natural substrate naphthalene (NDO) as well as with ethylbenzene (BDO).<sup>153,173</sup> Since two different protocols for the heterologous expression of NDO have been described, both literature conditions were tested which differ in the medium, expression temperature and time as well as in induction.<sup>145,174</sup> Applying the protocol described by Parales and coworkers, cells were grown in minimal medium and after 2 h of IPTG-induced protein expression at 30°C, full conversion with naphthalene as substrate could be achieved in whole cell biotransformations.<sup>145</sup> In contrast, activities were significantly lower applying the conditions reported by Seo *et al.* with cell growth over night at 37°C to induce protein expression by autoinduction in LB medium.<sup>174</sup> Due to higher product formations, expression of NDO was hence performed as described by the group of Parales.

For CDO, the plasmid pIP107D containing all RO components according to literature<sup>147</sup> lacked about 400 bp at the C-terminal end of the reductase *cumA4*. Thus, the reductase gene was restored

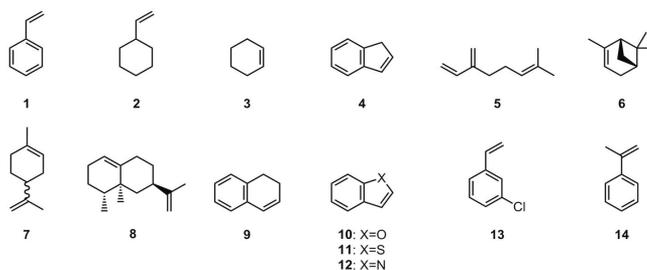
using plasmid pIP103 as template for PCR which harbored the corresponding *cis*-dihydrodiol dehydrogenase and the extradiol dioxygenase next to the RO genes.<sup>147</sup> By PCR, an *EcoRI* restriction site was appended after the stop codon of *cumA4* and the PCR product was used for the assembly of plasmid pIP107D by restriction digest (*SaII*, *EcoRI*) and ligation. Reconstruction of the reductase as well as functional expression of all RO components was confirmed in biotransformations with the natural substrate cumene. In contrast to NDO and BDO, the expression protocol for CDO was slightly varied from literature conditions.<sup>73</sup> Instead of 37°C, the expression temperature was lowered to 30°C while higher IPTG concentrations of 200 µM instead of 10 µM were used for induction. With the adapted conditions, a stronger blue color formation was observed with indole as substrate.

### 3.1.2. Pre-screening of different olefins

In order to elucidate the substrate spectrum of RO enzymes, an initial screening of various olefins was performed to define a panel of different compounds for the further characterization of NDO, BDO and CDO. For alkene selection, special attention was paid to include a variety of structurally diverse compounds as well as different alkene types (mono-, *gem*-di-, *cis*-di- and tri-substituted olefins). The panel of olefins that was examined during the pre-screening process comprised arene-substituted alkenes like styrene **1** and its derivatives 3-chlorostyrene **13** and  $\alpha$ -methylstyrene **14** as well as various terpenes including the monoterpenes myrcene **5**, (+)- $\alpha$ -pinene **6**, (*R*)- and (*S*)-limonene **7** and the sesquiterpene (+)-valencene **8**. In addition, indene **4** and other chromane substrates like 1,2-dihydronaphthalene **9**, 2,3-benzofuran **10**, 1-benzothiophene **11** and indole **12** were examined as well as the linear and cyclic alkenes vinylcyclohexane **2** and cyclohexene **3**, both possessing an unconjugated C=C double bond (Scheme 16).

Of the 14 olefins that were assayed, most compounds were accepted as substrates by BDO, CDO or NDO. The only exceptions comprised terpenes **6** and **8** which possess a sterically demanding structure when compared to the planar nature of the natural aromatic compounds and no conversion was detected with the examined ROs (Table 11). Next to **6** and **8**, the three enzymes showed differences in their substrate scope for the remaining olefins. In contrast to NDO or CDO, BDO possessed only minor activities < 10% for alkenes **2**, **5** and **7** while no conversion could be detected for compound **13** (Table 11), reducing its substrate scope compared to the other enzymes. Yet, also CDO showed low activities with < 10% conversion for alkene **13**.

According to detected masses in GC-MS analysis, monohydroxylated products were obtained next to the corresponding diols for alkenes **3**, **4**, **5**, **7** and **14** depending on the RO enzyme. Furthermore, for olefin **9**, desaturation of the substrate as already described for NDO was also observed with BDO and CDO.<sup>175</sup> Even though containing different positions for RO-catalyzed oxidations, compounds **1**, **2** and **7** were converted with good regioselectivities in biotransformations with NDO or CDO. For the other alkenes of the panel, multiple products were obtained with the main product constituting  $\leq 90\%$  of the mixture. As the three ROs differed in their regioselectivities, substrates **1**, **10**, **11**, **13** and **14** yielded diverse products but even when the same products were formed, ratios with NDO, BDO and CDO always varied.



**Scheme 16:** Substrate panel of alkenes examined in initial *in vivo* biotransformations with BDO, CDO and NDO.

**Table 11:** Activity of NDO, BDO and CDO towards selected olefins (10 mM) in whole cell biotransformations.

Olefin	1	2	3	4	5	6	( <i>R</i> )-7 <sup>[a]</sup>	8	9	10	11	12	13	14
NDO	+++	+	+	+++	+	-	n.d.	-	+++	++	++	++	+++	+++
BDO	+++	(+)	++	+++	(+)	-	(+)	-	+	+	+	++	-	++
CDO	+++	++	+	+++	++	-	++	-	+	+++	+++	+	(+)	+++

+++ : Conversion of 70- > 99%, ++ : Conversion of 40-70%, + : Conversion of 10-40%, (+) : < 10% conversion, -: no conversion detected, n.d.: not determined. Conversion was estimated from the GC areas of substrate and product. As no derivatization of biotransformation products prior to GC-analysis was performed in the pre-screening, very low product formations might have been missed due to the higher detection limit of underivatized compounds. <sup>[a]</sup> Conversion was only determined for (*R*)-7 as the (*S*)-enantiomer was not commercially available in optically pure form containing also (*R*)-7 (68% *ee*). With both enantiomers, a monohydroxylated product was obtained according to GC-MS analysis and product formation with (*S*)-7 as substrate was observed with CDO. Reactions in the pre-screening were performed as duplicates.

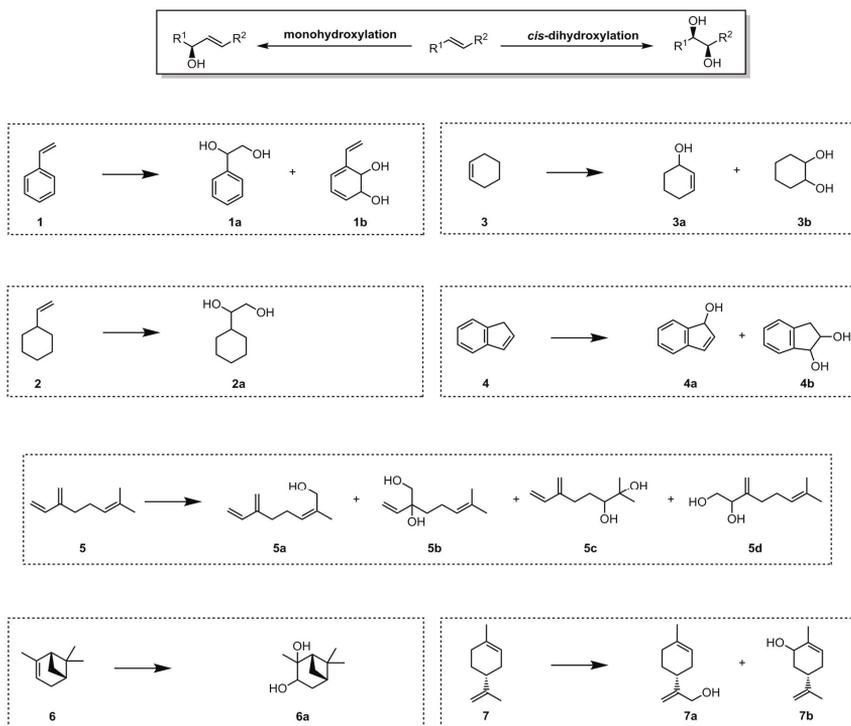
In order to determine product formations as well as stereo- and regioselectivities of BDO, CDO and NDO for different olefins, the original substrate panel of 14 alkenes was reduced in size. With regard to the results from the pre-screening, alkenes 1-7 were selected for further investigation. Herein, compounds 1, 2 and 7 were chosen due to the good regioselectivities obtained for CDO or NDO. While 1 and 2 allowed evaluating the dihydroxylation of conjugated *versus* unconjugated linear alkenes, biotransformations with terpene 7 solely yielded the alcohol product instead of the diol. Furthermore, the cyclic alkene 3 was included in the panel since it represents the smallest olefin in size of the selected compounds. Of the chromane substrates, alkene 4 was selected for further examination as conversions > 70% were obtained for all enzymes, not having been described in literature for CDO and BDO. In addition, the alkene-1,2-diol represents a potential precursor for Indinavir which is applied as HIV inhibitor.<sup>9</sup> With polyene 5, a tri-substituted C=C double bond was included while combining three different alkene types in one molecule. Also terpene 6 contained a tri-substituted alkene bond but was not accepted by the wild type enzymes. Yet, with regard to the second part of this work, the optimization of wild type ROs for olefin oxyfunctionalization, biotransformation of 6 was further examined, too.

Even though limited to seven compounds, this panel of structurally diverse olefins included arene-substituted, cyclic and linear alkenes comprising mono- (1, 2, 5), *gem*-di- (5, 7), *cis*-di- (3, 4) and tri-

substituted C=C double bonds (**5**, **6**, **7**), enabling the evaluation of the RO-catalyzed AD of different alkene types against heavy metal catalysts like OsO<sub>4</sub>. Furthermore, several of the substrates allowed oxidation at different positions as they possess either an aromatic ring besides the alkene moiety (**1** and **4**) or include several aliphatic C=C double bonds (**5** and **7**), hence representing interesting substrates to compare ROs for differences in regioselectivity.

### 3.1.3. Characterization of ROs towards a panel of selected alkene substrates

Conversions of alkenes **1-7** selected in the pre-screening process were characterized in *in vivo* biotransformations with 10 mM substrate and 0.2 g mL<sup>-1</sup> cells expressing BDO, CDO or NDO. For identification of biotransformation products, the reaction set-up was either up-scaled in order to isolate the compounds of interest and to determine their structure by NMR analysis or products were identified by comparison with an authentic standard (Scheme 17). For product identification and quantification as well as determination of stereoselectivities, (*R*)- and (*S*)-**2a**, (*R*)-(+)-**3a**, (1*R*,2*S*)-**4b**, (*E*)-**5a**, *rac*-**5b**, (*R*)- and (*S*)-**5c**, *rac*-**5d**, (1*S*,2*S*,3*R*,5*S*)-(+)-**6a**, **7a** as well as (1*S*,5*S*)- and (1*R*,5*R*)-**7b** were synthesized as described in the experimental section.



**Scheme 17:** Selected substrate panel of alkenes for the characterization of *in vivo* biotransformations with BDO, CDO and NDO as well as detected biotransformation products.

**Table 12:** Product formations and distributions for CDO, NDO and BDO with alkenes 1-7 (10 mM).

Olefin	Product formation [%] <sup>[a]</sup>			Product distribution [%] <sup>[b]</sup>		
	CDO	NDO	BDO	CDO	NDO	BDO
<b>1</b>	71 ± 6	> 99	90 ± 7	0.3:99.7 ( <b>1a:1b</b> )	> 99 ( <b>1a</b> )	59:41 ( <b>1a:1b</b> )
<b>2</b>	44 ± 6	20 ± 3	< 1	>95 ( <b>2a</b> )	> 95 ( <b>2a</b> )	n.d.
<b>3<sup>[b]</sup></b>	36 ± 4	27 ± 5	59 ± 6	80:20 ( <b>3a:3b</b> )	51:49 ( <b>3a:3b</b> )	76:24 ( <b>3a:3b</b> )
<b>4<sup>[c]</sup></b>	80 ± 7	87 ± 10	> 99	64:36 ( <b>4a:4b</b> )	38:62 ( <b>4a:4b</b> )	10:90 ( <b>4a:4b</b> )
<b>5</b>	29 ± 6	34 ± 11	5 ± 0.1	48:18:18:17 ( <b>5a:5b:5c:5d</b> )	12:41:3:44 ( <b>5a:5b:5c:5d</b> )	56:14:14:15 ( <b>5a:5b:5c:5d</b> )
<b>6</b>	-	-	-	-	-	-
( <i>R</i> )- <b>7<sup>[d]</sup></b>	46 ± 10	n.d.	< 1	>95 ( <b>7b:7a</b> )	n.d.	n.d.

<sup>[a]</sup> Product formations were determined by GC-FID analysis after 24 h using standard curves of the respective products. Standard deviations also include variations in the expression level as triplicates from at least two different expression batches were used for the determination of product formations. <sup>[b]</sup> Product distributions were determined by GC-FID. Cyclohexanone formation resulting from isomerization of **3a** was added to the total amount of **3a**. <sup>[c]</sup> Traces of 1-indanol (1-2%) were detected resulting from the RO-catalyzed oxidation of indane trace amounts present in **4**. <sup>[d]</sup> Since (*S*)-**7** was only available as mixture of both substrate enantiomers (68% *ee*), product formations were solely determined with (*R*)-**7** being commercially available in enantiopure form. -: not detected. n.d.: not determined.

**Table 13:** Stereoselectivities for CDO, NDO and BDO.

Product	Stereoselectivity <i>ee</i> or <i>de</i> [%] <sup>[a]</sup>		
	CDO	NDO	BDO
<b>1a</b>	43 ± 3 ( <i>R</i> - <b>1a</b> )	80 ± 1 ( <i>R</i> - <b>1a</b> )	8 ± 2 ( <i>R</i> - <b>1a</b> )
<b>2a</b>	91 ± 3 ( <i>R</i> - <b>2a</b> )	> 98 ( <i>R</i> - <b>2a</b> )	n.d.
<b>3a</b>	74 ± 0.5 ( <i>R</i> - <b>3a</b> )	65 ± 2 ( <i>R</i> - <b>3a</b> )	90 ± 1 ( <i>R</i> - <b>3a</b> )
<b>4a</b>	10 ± 1 ( <i>R</i> - <b>4a</b> )	71 ± 4 ( <i>S</i> - <b>4a</b> )	45 ± 1 ( <i>S</i> - <b>4a</b> )
<b>4b</b>	38 ± 1 (1 <i>S</i> ,2 <i>R</i> - <b>4b</b> )	76 ± 1 (1 <i>R</i> ,2 <i>S</i> - <b>4b</b> )	42 ± 1 (1 <i>R</i> ,2 <i>S</i> - <b>4b</b> )
<b>5b</b>	78 (n.d., <b>5b</b> )	> 95 (n.d., <b>5b</b> )	n.d.
<b>5c</b>	> 95 ( <i>R</i> - <b>5c</b> )	n.d.	-
<b>5d</b>	47 (n.d., <b>5d</b> )	> 95 (n.d., <b>5d</b> )	n.d.
<b>6a</b>	-	-	-
( <i>R</i> )- <b>7</b>	> 98 (1 <i>R</i> ,5 <i>S</i> - <b>7b</b> )	n.d.	n.d.
( <i>S</i> )- <b>7</b>	95 (1 <i>R</i> ,5 <i>R</i> - <b>7b</b> )	n.d.	n.d.

<sup>[a]</sup> Enantiomeric and diastereomeric excess values were determined by GC-FID or HPLC analysis using a chiral stationary phase. For chiral GC and HPLC chromatograms see supplementary. To determine standard deviations, triplicates from at least two different expression batches were used. -: not detected. n.d.: not determined.

Of the seven selected model substrates, compound **1** was accepted by each of the three enzymes. Depending if oxidation took place at the aromatic ring or the terminal vinyl group of the olefin, the AD of **1** yielded either alkene-1,2-diol **1a** or arene-1,2-dihydrodiol **1b** (Scheme 17). Selective dihydroxylation of **1** by NDO exclusively at the terminal vinyl group has been already reported in literature.<sup>169</sup> In contrast, during CDO- and BDO-catalyzed biotransformations, **1b** was formed besides **1a**. While CDO showed an excellent regioselectivity for the aromatic moiety of the substrate (0.3:99.7 **1a:1b**), a mix of **1a** and **1b** (59:41) was observed for biotransformations with BDO. All enzymes had a preference for the (*R*)-enantiomer of **1a** with NDO giving the best stereoselectivity (80% *ee*). Yielding *ee* values of only 8% and 43% for (*R*)-**1a**, stereoselectivities obtained with BDO and CDO were low (Table 13).

Even though it has been reported in literature that non-conjugated linear alkenes are generally found to be poor substrates for *cis*-dihydroxylation by ROs,<sup>75</sup> CDO catalyzed the oxidation of **2** with 44% product formation and good stereoselectivity (*ee* = 91% for diol (*R*)-**2a**). Compared to CDO, biotransformations with NDO gave even better *ee* values of > 98% for (*R*)-**2a**, albeit conversions were low (20%, Table 12). For the cyclic alkene **3**, the second non-conjugated olefin in this study, all examined ROs catalyzed the allylic hydroxylation of the substrate yielding **3a** besides diol **3b** (Scheme 17). Thereby, enzymes displayed varying product distributions with NDO forming the highest ratio of diol **3b** (49%) whereas alcohol **3a** was the main product with CDO and BDO (80% and 76%, respectively, Table 12).

The arene-substituted olefin **4** was converted by NDO, CDO and BDO yielding diol **4b** and the monohydroxylated product **4a** (Scheme 17). No oxyfunctionalization of the aromatic ring was found to take place with CDO and BDO as observed for compound **1**. Furthermore, BDO, NDO and CDO showed only low to moderate stereoselectivities for the conversion of **4**. Interestingly, CDO favored opposite enantiomers for both **4a** and **4b** compared to BDO and NDO albeit the obtained stereoselectivities were low. CDO-catalyzed biotransformation of **4** gave (*R*)-**4a** with an *ee* value of 10% whereas BDO and NDO yielded (*S*)-**4a** with significantly higher stereoselectivities (45% and 71% *ee*, respectively). Also for product **4b**, NDO showed the best stereoselectivity with 76% *ee* for the (1*R*,2*S*)-enantiomer. Like for alcohol **4a**, CDO gave the opposite enantiomer of **4b** compared to BDO and NDO. With 42% and 38% *ee*, the stereoselectivities for the diol were however low for BDO and CDO. Yet, moderate stereoselectivities are also obtained for the OsO<sub>4</sub>-catalyzed AD of **4** as *cis*-di-substituted alkenes represent challenging substrates for this reaction.<sup>176,177</sup> Furthermore, the product ratios of **4a** and **4b** differed between the three enzymes. Forming 90% of diol **4b**, BDO showed the best regioselectivity. Likewise, NDO-catalyzed biotransformation of **4** gave 62% **4b** whereas with CDO, alcohol **4a** was the main product formed (64%, Table 12).

Even though applied as flavors and fragrances,<sup>178</sup> the oxyfunctionalization of terpenes by RO enzymes has rarely been described. Hence, compounds **5**, **6** and **7** were tested as substrates for the RO-catalyzed oxidation. Biotransformations of the linear polyalkene **5** with NDO, BDO and CDO gave a mixture of **5a**, **5b**, **5c** and **5d**. In contrast to the SeO<sub>2</sub>-mediated allylic oxidation of **5** yielding the (*E*)-isomer of **5a**,<sup>161</sup> (*Z*)-**5a** was exclusively formed during biotransformations. For CDO and BDO, alcohol (*Z*)-**5a** was the main product (48% and 56%, respectively) of the oxidation of **5**. With NDO, diols **5b** and **5c** were obtained in a ratio of 41:44 while the reaction proceeded with excellent

stereoselectivities > 95% for both products. With CDO, high stereoselectivities > 95% were obtained for diol **5c**, however, *ee* values for **5b** and **5d** were only moderate (78% and 47% *ee*, respectively). Nevertheless, having the ability to target the electron-deficient C=C double bonds of **5** is of special interest as in Sharpless AD, osmylation of unsymmetrical polyenes preferentially occurs at the most electron-rich double bond<sup>7,11</sup> and during OsO<sub>4</sub>-catalyzed oxidation of **5**, diol **5c** was formed selectively. As diols **5b** and **5d** had hence to be synthesized with KMnO<sub>2</sub> giving a racemic mixture of both compounds,<sup>162</sup> the absolute configurations in RO-catalyzed biotransformations could not be assigned. Yet, both CDO and NDO showed the same stereopreference. Due to the low product formations with BDO (5%), stereoselectivities for diols **5b**, **5c**, and **5d** could not be determined. For CDO-catalyzed biotransformations of limonene **7**, allylic hydroxylation yielded alcohol **7b** as main product. Depending if the (*R*)- or the (*S*)-enantiomer was employed as a substrate, (1*R*,5*S*)-(+)- or (1*R*,5*R*)-(-)-**7b** was formed with an *ee* value of > 98% and 95%, respectively. With BDO, conversions were < 1% and hence, products were not further characterized (Table 12). As already mentioned in the pre-screening, terpene **6** with its sterically demanding structure and tri-substituted alkene bond could not be converted by BDO, NDO or CDO. Accordingly, compound **6** also proved to be a challenging substrate for Sharpless AD under standard conditions requiring elevated temperatures ≥ 70°C.<sup>163</sup>

### 3.2. Rational design – Optimization of ROs for selective olefin oxyfunctionalization

From the selected substrate panel, CDO, NDO and BDO converted six of the seven alkenes. Yet, unlike described for their natural substrates, enzymes displayed only low conversions (< 50%) and modest stereoselectivities (≤ 80% *ee*) for several of the tested olefins. In particular, the sterically demanding terpene **6** proved to be a challenging substrate for the RO-catalyzed oxyfunctionalization and no conversion could be detected for NDO, BDO or CDO. High stereoselectivities ≥ 90% were only obtained for diols **2a** (NDO and CDO), **5c** (CDO), **5b** and **5d** (NDO), alcohol **7b** (CDO) as well as for the BDO-catalyzed formation of alcohol **3a** (Table 12 and Table 13). The results are in accordance with literature where it has been shown that the application of ROs in organic synthesis is partly restricted by their poor stereoselectivity and activity, especially for linear alkenes.<sup>74,75</sup> To overcome these limitations, changes in the active site topology of RO catalysts were introduced to generate variants with improved characteristics for the selective oxyfunctionalization of olefins in order to optimize these enzymes for the conversion of unnatural substrates.

#### 3.2.1. Selection of amino acid residues by phylogenetic comparison

Since the focus of this project has been placed on the selective oxyfunctionalization of aliphatic C=C double bonds, the identification of amino acid residues influencing selectivity and activity of ROs was one of its key issues. Although closely related (46-78% similarity of the α-subunits containing the active site)<sup>c</sup>, NDO, BDO and CDO displayed a strong bias with respect to substrate

<sup>c</sup> The sequence similarity was determined by performing pairwise sequence alignments with the European Molecular Biology Open Software Suite (EMBOSS) applying the Needleman-Wunsch algorithm.<sup>221</sup>

specificity as well as selectivity. Caused by variations in the active site topology of the three enzymes that influence enzyme characteristics, selectivities and activities for the different alkene types were strongly dependent on the RO employed.

In order to identify amino acid residues that are essential for RO selectivity, a structural as well as sequence alignment of the active sites of CDO, BDO and NDO was performed (Figure 10 and Figure S30). Examining all amino acid residues within 8 Å of the catalytic iron center, only those side chains were considered for mutagenesis that differed between ROs in structurally equivalent positions. Using this strategy, conserved amino acids which are essential for activity like the 2-His-1-carboxylate facial triad coordinating the catalytic iron were excluded. Furthermore, amino acid side chains pointing away from the active site were not taken into account as these residues might most likely not contribute to the orientation of the substrate in the binding pocket.

**Table 14:** Amino acid residues within 8 Å of the catalytic iron center of NDO, CDO, BDO and TDO.

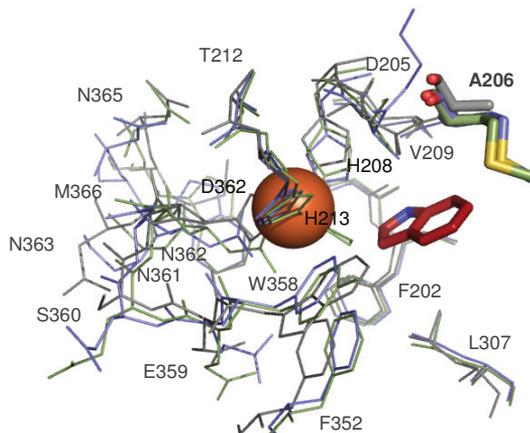
Position <sup>[a]</sup>	NDO	CDO	BDO	TDO	Position <sup>[a]</sup>	NDO	CDO	BDO	TDO
<b>201</b>	<b>N</b>	<b>Q</b>	<b>Q</b>	<b>Q</b>	307	L	L	L	L
202	F	F	F	F	352	F	F	F	F
205	D	D	D	D	<b>358</b>	<b>W</b>	<b>Y</b>	<b>F</b>	<b>F</b>
<b>206</b>	<b>A</b>	<b>M</b>	<b>M</b>	<b>M</b>	359	E	E	E	E
208*	H	H	H	H	361	D	D	D	D
<b>209</b>	<b>V</b>	<b>A</b>	<b>A</b>	<b>A</b>	362*	D	D	D	D
<b>212</b>	<b>T</b>	<b>A</b>	<b>A</b>	<b>S</b>	365	N	N	N	N
213*	H	H	H	H	<b>366</b>	<b>M</b>	<b>W</b>	<b>W</b>	<b>W</b>

<sup>[a]</sup> Numbering refers to NDO from *Pseudomonas* sp. NCIB 9816-4. \* Amino acids constituting the 2-His-1-carboxylate facial triad. Amino acid residues differing between the enzymes are marked in bold.

Of the 16 amino acids located within 8 Å of the mononuclear iron center, positions 201, 206, 209, 212, 358 and 366 (numbering refers to NDO) varied between the enzymes (Table 14). Amongst these 6 residues, the methionine side chain at position 220 and 232, respectively, in BDO and CDO was located closest to the substrate (Figure 10, residues correspond to position 206 in NDO). With regard to the reported interactions of methionine with aromatic systems,<sup>179,180</sup> residues 232 and 220 might be involved in the positioning of aryl-substituted substrates like **1** and **4** in the active sites of both ROs, representing two of the most striking examples for differing selectivities between the three enzymes. Interestingly, the methionine residue pointing into the active site was not only present in CDO and BDO, but also at the corresponding position of TDO from *P. putida* F1 (M220, Table 14).<sup>181</sup> As these enzymes all catalyzed the dihydroxylation of compound **1** at the aromatic ring, this was a further indication for the side chain of methionine playing an important role in the orientation of **1** in the substrate binding pocket.

Mutations at the corresponding position 206 in NDO have been already reported to show a significant effect on product formations as well as selectivities for several aromatic compounds.<sup>125</sup> Furthermore, variation of M220 in TDO resulted in an expanded substrate spectrum towards 1,2,4,5-tetrachlorobenzene,<sup>182</sup> suggesting an influence of this position on RO specificity. In contrast, only slight alterations in product formation were observed for variations of N201, W358 or M366 in

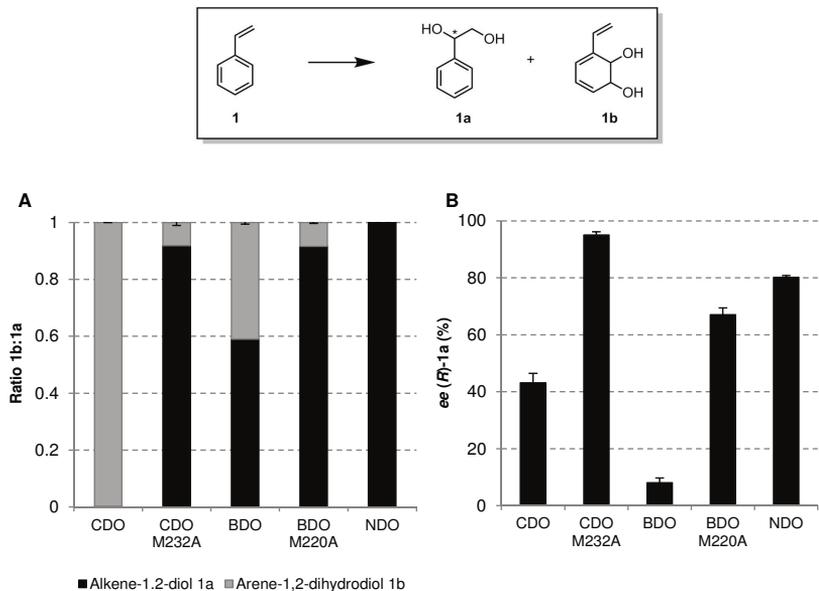
NDO<sup>145</sup> while no mutational studies of the remaining variable residues have been published for the selected RO enzymes. Considering these data, methionine residues 220 and 232, respectively, in BDO and CDO were selected amongst the variable positions and exchanged by the corresponding alanine found in NDO. Of the three ROs, NDO not only catalyzed dihydroxylation of **1** exclusively at the terminal vinyl group but also gave the best stereoselectivities for most of the tested substrates (Table 13).



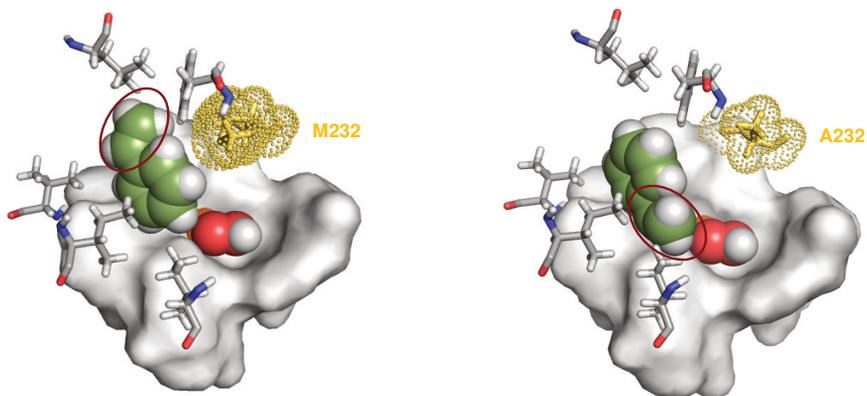
**Figure 10:** Structural alignment of amino acid residues in a distance of 8 Å around the catalytic active iron (orange sphere) from CDO (PDB code: 1WQL, green lines),<sup>170</sup> NDO (PDB code: 1O7N, grey lines)<sup>101</sup> and BDO (blue lines). The alignment was performed using the PyMOL molecular graphics system, version 1.2r1 Schrödinger, LLC. For BDO, a homology model was created based on the TDO crystal structure (PDB code: 3EN1)<sup>171</sup> using SWISS-MODEL.<sup>183</sup> Residues M232 (CDO, green), A206 (NDO, grey) and M220 (BDO, blue) chosen for mutagenesis are shown as sticks. The bound substrate (indole) is shown as red sticks. Positions refer to numbering in NDO from *Pseudomonas* sp. NCIB 9816-4.

### 3.2.2. Rational design - Characterization of generated RO variants

After altering their active site topology by site-directed mutagenesis, CDO M232A and BDO M220A were characterized with the selected alkene panel. For conversion of **1**, both RO variants showed indeed a dramatic shift in regioselectivity towards the terminal vinyl group and product formation of alkene-1,2-diol **1a** was increased to 92% compared to 0.3% and 59% for CDO and BDO wild type enzymes (Figure 11A). This change in regioselectivity was accompanied by a significant increase in stereospecificity (Figure 11B). For CDO M232A, the *ee* for (*R*)-**1a** improved from 43% to 95% which is even better than *ee* values obtained with NDO (80% *ee*) and represents one of the highest stereoselectivities reported so far for the dihydroxylation of a linear aliphatic substrate by ROs. Moreover, the single amino acid substitution of methionine to alanine also enhanced the stereoselectivity of BDO M220A for **1a** by about 8-fold to 67%. Yet, these changes in the catalytic site topology did not cause a dramatic change in activity. For BDO M220A, product formation decreased only slightly from 90% for the wild type enzyme to 81% whereas for CDO M232A an increase to 97% product formation was observed (Table 16).



**Figure 11:** (A) Ratios of products **1a** (black) and **1b** (grey) for the biotransformation of styrene **1** with CDO, CDO M232A, BDO, BDO M220A and NDO. (B) ee Values for alkene-1,2-diol **1a** with CDO, CDO M232A, BDO, BDO M220A and NDO.



**Figure 12:** Docking of styrene **1** (green spheres) in the active site of CDO wild type (left) and CDO M232A (right). Methionine at position 232 in the wild type and alanine inserted by site-directed mutagenesis in the CDO variant are shown as yellow dots. Surrounding amino acids that constitute the substrate binding pocket are shown in grey (surface and sticks). In the crystal structure (PDB code: 1WQL),<sup>170</sup> O<sub>2</sub> (red spheres) is bound to the catalytic non-heme iron (orange spheres, only partly visible, located behind the red spheres). The C=C double bond of the substrate is circled in red. The homology model of the CDO variant M232A was created with SWISS-MODEL<sup>183</sup> based on the CDO crystal structure (PDB code: 1WQL).<sup>170</sup>

Regarding these differences in selectivities of wild type enzymes and variants, docking of **1** into the active site of CDO and CDO M232A was performed. Depending on the shape of the chiral binding pocket, either the aromatic ring (CDO wild type) or the vinyl group (CDO M232A) of **1** is oriented close to the catalytic iron center (Figure 12), supporting the hypothesis of different substrate positions in the active sites of wild type and mutant enzymes.

With respect to the dramatic changes in regio- and stereoselectivity that can be introduced by the single point mutation in CDO and BDO, the effect of methionine at the corresponding position A206 in NDO was investigated for conversion of **1**. However, the diminishment of the active site volume by the introduction of the bulky methionine side chain caused a significant decrease in activity to only 20% product formation (Table 15). In contrast to CDO and BDO, the amino acid exchange did not induce an alteration in regioselectivity. Yet, a slight increase in *ee* for alkene-1,2-diol **1a** was observed for NDO A206M, suggesting that also in NDO, this position has an influence on selectivity for olefin conversion (Table 15). As the introduction of methionine into the substrate binding pocket caused a substantial loss of activity, mutant NDO A206I was generated to investigate the effect, a residue smaller than methionine but larger than alanine might have on oxidation of **1**. Furthermore, this variant has already been described in literature to influence enzyme selectivity during arene oxidation.<sup>125</sup> Whereas the decrease in activity for NDO A206I (80% product formation) was less severe than with NDO A206M, this mutation neither showed a significant impact on stereoselectivity nor on regioselectivity for substrate **1** (Table 15).

**Table 15:** Product formations as well as regio- and stereoselectivities for the conversion of **1** (10 mM) with NDO wt and NDO variants A206M and A206I.

RO	Product formation [%] <sup>[a]</sup>	Regioselectivity [%]	Stereoselectivity <i>ee</i> [%] <sup>[b]</sup>
NDO wt	> 99	> 99 ( <b>1a</b> )	80 ± 1 ( <i>R</i> - <b>1a</b> )
NDO A206M	20 ± 1	> 99 ( <b>1a</b> )	85 ± 1 ( <i>R</i> - <b>1a</b> )
NDO A206I	80 ± 11	> 99 ( <b>1a</b> )	81 ± 3 ( <i>R</i> - <b>1a</b> )

<sup>[a]</sup> Product formations were determined by GC-FID analysis after 24 h using standard curves of the respective products. Standard deviations for NDO variants were determined with triplicates of one expression batch. <sup>[b]</sup> Enantiomeric excess values were determined by GC-FID or HPLC analysis using a chiral stationary phase.

Besides alkene **1**, substrates **2-7** were re-examined with CDO M232A and BDO M220A variants. Compared to the wild type enzymes, mutants displayed not only altered selectivities but also changes in product formations for several of the investigated compounds. Especially the use of CDO M232A resulted in enhanced stereoselectivities for compounds displaying higher *ee* values with NDO than CDO wild type. For the unconjugated linear olefin **2**, CDO M232A showed an enhanced conversion and stereoselectivity compared to the wild type enzyme. In biotransformations with CDO M232A, the conversion could be increased about 2-fold, resulting in 93% overall product formation (Table 16). Of the two possible enantiomers, (*R*)-**2a** was formed with an excellent stereoselectivity of > 98% *ee* instead of 91% *ee* for the wild type enzyme. In contrast, for BDO M220A, exchange of methionine by alanine did not have a significant effect on product formation and conversion of **2** was still below < 1% as already observed for BDO wild type.

Unlike for **1** and **2**, the *ee* values for the oxidation of the second non-conjugated alkene **3** decreased in biotransformations with both variants. BDO M220A showed a slight decrease in stereoselectivity from 90% to 86% *ee* whereas for CDO M232A, the *ee* value dropped from 74% to 28%. Despite this drastic change in stereoselectivity, product formation of CDO M232A was not affected (Table 12 and Table 16). Yet, for BDO M220A, conversion decreased from 59% to 5%.

**Table 16:** Product formation, distribution and stereoselectivities for CDO M232A and BDO M220A with substrates 1-(*R*)-**7** (10 mM). Values determined for the wild type enzymes are shown in Table 12 and Table 13.

Olefin	Product formation [%] <sup>[a]</sup>		Product distribution [%]		Stereoselectivity <i>ee</i> or <i>de</i> [%] <sup>[b]</sup>	
	CDO	BDO	CDO	BDO	CDO	BDO
	M232A	M220A	M232A	M220A	M232A	M220A
<b>1</b>	97 ± 10	81 ± 10	92:8 ( <b>1a:1b</b> )	92:8 ( <b>1a:1b</b> )	95 ± 1 ( <i>R</i> - <b>1a</b> )	67 ± 2 ( <i>R</i> - <b>1a</b> )
<b>2</b>	93 ± 5	< 1	> 95 ( <b>2a</b> )	n.d.	> 98 ( <i>R</i> - <b>2a</b> )	n.d.
<b>3</b> <sup>[c]</sup>	36 ± 4	5 ± 1	82:18 ( <b>3a:3b</b> )	68:32 ( <b>3a:3b</b> )	28 ± 1 ( <i>R</i> - <b>3a</b> )	86 ± 1 ( <i>R</i> - <b>3a</b> )
<b>4</b> <sup>[d]</sup>	80 ± 11	82 ± 13	56:44 ( <b>4a:4b</b> )	22:78 ( <b>4a:4b</b> )	87 ± 1 ( <i>S</i> - <b>4a</b> ) 54 ± 4 ( <i>1R,2S</i> - <b>4b</b> )	82 ± 1 ( <i>S</i> - <b>4a</b> ) 33 ± 1 ( <i>1R,2S</i> - <b>4b</b> )
<b>5</b>	> 99	37 ± 4	71:3:21:4 ( <b>5a:5b:5c:5d</b> )	31:39:15:15 ( <b>5a:5b:5c:5d</b> )	> 95 (n.d. <b>5b</b> ) > 95 ( <i>R</i> - <b>5c</b> ) n.d.	n.d.
<b>6</b>	33 ± 7	-	> 95% ( <b>6a</b> )	-	> 95 ( <i>1S,2S,3R,5S</i> - <b>6a</b> )	-
( <i>R</i> )- <b>7</b>	> 99	< 1	95:5 ( <b>7b:7a</b> )	n.d.	> 98 ( <i>1R,5S</i> - <b>7b</b> )	n.d.

<sup>[a]</sup> Product formations were determined by GC-FID analysis after 24 h using standard curves of the respective products. Standard deviations also include variations in the expression level as triplicates from at least two different expression batches were used for the determination of product formations. <sup>[b]</sup> Enantiomeric and diastereomeric excess values were determined by GC-FID or HPLC analysis using a chiral stationary phase. For **6a**, diastereomeric excess values were in addition verified by <sup>1</sup>H NMR spectroscopy as only one of the two possible diastereomers was available as standard. For chiral GC or HPLC chromatograms and NMR spectra see supplementary. <sup>[c]</sup> Cyclohexanone formation resulting from isomerization of **3a** was added to the total amount of **3a**. <sup>[d]</sup> Traces of 1-indanol (1-2%) were detected resulting from the RO-catalyzed oxidation of indane trace amounts present in **4**. -: not detected. n.d.: not determined.

With the arene-substituted alkene **4**, the most pronounced change in stereoselectivity could be observed for CDO M232A. The single point mutation in the active site of CDO enhanced not only the *ee* value by about 9-fold but also led to a change in the absolute configuration of the product **4a**, yielding the (*S*)-enantiomer with 87% *ee*. Likewise, for diol **4b**, the increase in *ee* from 38% to 54% was accompanied by a switch from (*1S,2R*)-**4b** to (*1R,2S*)-**4b**, turning CDO and CDO M232A into enantiocomplementary enzymes for the conversion of **4**. An altered stereoselectivity was also observed for BDO variant M220A. Yet, whereas the *ee* for (*S*)-**4a** was improved from 45% to 82%, the stereoselectivity for the diol **4b** decreased to 33% (Table 13 and Table 16).

With CDO M232A and BDO M220A, enhanced substrate oxidation up to 7-fold was achieved for the linear polyalkene **5**. In CDO M232A-catalyzed biotransformations, the stereoselectivity for product **5b** was increased to > 95% while maintaining the excellent stereoselectivity for diol **5c** already observed for the wild type enzyme (Table 13 and Table 16). Additionally, both variants exhibited altered product ratios of (*Z*)-**5a**, **5b**, **5c** and **5d** with CDO M232A forming alcohol (*Z*)-**5a** as main product with 71% (Table 16).

Like the wild type enzyme, CDO M232A-catalyzed oxidation of (*R*)-**7** gave alcohol **7b** with good regioselectivity, yet, conversion could be improved 2-fold from 46% to > 99%. Furthermore, for oxidation of (*S*)-**7**, the *de* slightly increased from 95% to > 98%. In contrast, for BDO M220A, the alteration of its active site topology did not influence conversion and product formation of **7b** was still < 1%.

For terpene **6**, the only compound of the panel not converted by the wild type enzymes, the increase in the size of the binding pocket of CDO M232A resulted in an extended substrate scope. Biotransformations with this variant yielded 33% product formation of diol **6a** with an excellent stereoselectivity of > 95% *de* for the (1*S*,2*S*,3*R*,5*S*)-(+)-diastereomer. However, besides position 232 (CDO) or 220 (BDO), the surrounding amino acids in the enzymes also seemed to control substrate specificity as with BDO M220A, still no conversion of **6** could be detected. Docking analysis suggested that oxyfunctionalization of the terpene in BDO M220A was sterically hindered as **6** could apparently not enter the active site and residues lining the entrance tunnel or the substrate binding pocket might prevent productive binding and therewith catalysis.

### 3.3. Site-saturation mutagenesis of CDO - Creation of a focused mutant library

In addition to rational design, site-saturation mutagenesis was applied for CDO to create a focused mutant library which was screened for variants displaying improved selectivities or activities for alkene oxyfunctionalization. In contrast to site-directed enzyme design, the effect of diverse amino acid residues at a distinct position could be examined using site-saturation mutagenesis. Having been identified as active site hotspot, methionine at position 232 in CDO was selected to be replaced with a balanced mix of 12 structurally different amino acid side chains. Furthermore, four additional positions (F278, I288, I336 and F378) located in the substrate binding pocket of CDO were assayed for their influence on selectivity and activity, mainly using small amino acid residues for mutagenesis to increase the active site volume.

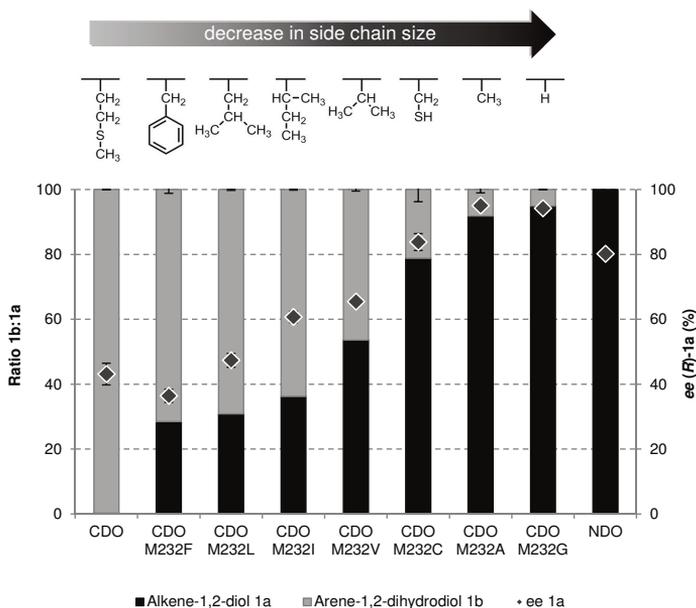
#### 3.3.1. Exploring the influence of various amino acid residues at position 232

Considering the strong impact of mutation M232A in CDO on selectivity, the effect of additional amino acid residues at position 232 was examined by site-saturation mutagenesis using the NDT codon degeneracy. NDT encodes for a mix of 12 polar, non-polar, charged, uncharged, aliphatic and aromatic amino acids (F, L, I, V, Y, H, N, D, C, S, R and G) which excludes structurally similar residues and thus reduces the screening effort compared to the NNK codon degeneracy.<sup>140</sup> 11 of these CDO variants (CDO M232F, L, I, V, Y, H, N, D, C, S and G) were tested for conversion of the

arene-substituted alkenes **1** and **4** as changes in regio- and stereoselectivity of CDO M232A were most pronounced with these substrates.

Of the generated mutants, only CDO variants M232Y, D and N lost complete activity towards compounds **1** and **4**. The other variants of the NDT library proved to be robust for the introduced mutations and formation of products **1a** and **1b** as well as **4a** and **4b** was detected even though with diminished activities compared to CDO wild type and CDO M232A (Table 17). For alkene **1**, higher product formations were achieved with all variants compared to the conversion of compound **4**. With substrate **4**, best activities were obtained for variants possessing sterically less demanding amino acid side chains (CDO M232C, S and G) but even with the most active mutants, product formations were still < 20% (Table 17).

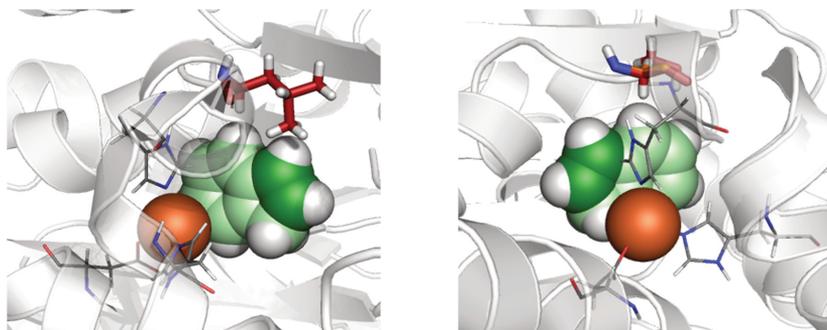
Despite this decrease in conversions, the variants of the NDT mutant library displayed interesting changes in selectivities which seemed to be correlated with the steric demand of the amino acid residue. For compound **1**, the amino acid side chain size at position 232 was shown to directly influence both regio- and stereoselectivity. In general, smaller amino acid residues favored the formation of alkene-1,2-diol **1a** accompanied by enhanced stereoselectivities (Figure 13). As an exception, the ratio of **1a** increased to 96% for CDO variant M232S, giving a higher regioselectivity for CDO M232A (92%) and CDO M232G (95%, Figure S31). Yet, in accordance with the general trend, both CDO M232A and CDO M232G showed better stereoselectivities for the alkene-1,2-diol (95% and 94% *ee* for (*R*)-**1a**) when compared to CDO M232S (88% *ee* for (*R*)-**1a**, Table 18).



**Figure 13:** Product distribution (bars) for the RO-catalyzed conversion of styrene **1** yielding alkene-1,2-diol **1a** (black) and arene-1,2-dihydrodiol **1b** (grey) as well as stereoselectivities for **1a** (squares).

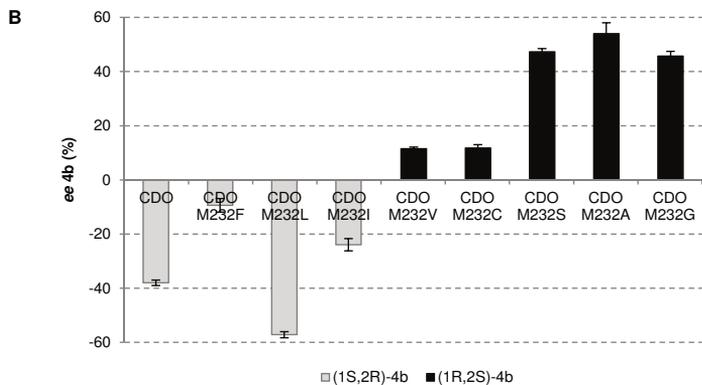
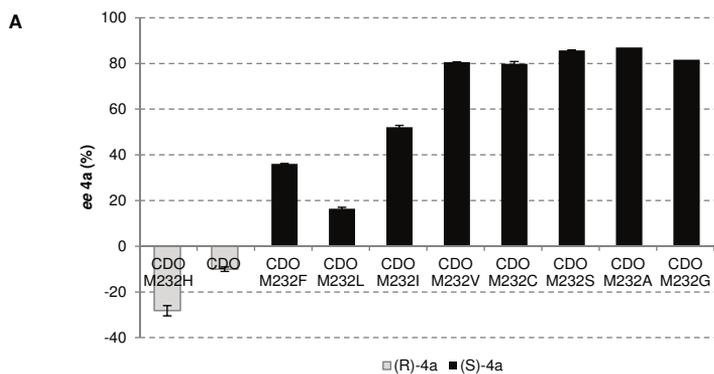
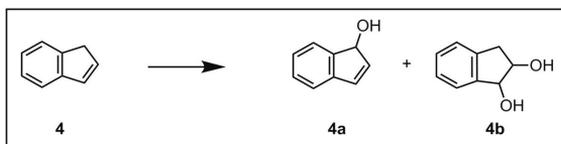
For alkene **4**, CDO M232H showed a 3-fold enhancement in *ee* to 28% (*R*)-**4a** whereas a switch in stereoselectivity from (*R*)-**4a** to (*S*)-**4a** was observed for the allylic alcohol with the other CDO variants compared to the wild type (Table 18). As demonstrated for **1a**, the stereoselectivity for (*S*)-**4a** generally increased with smaller amino acid residues (Figure 15A). For the diol formation, large side chains (F, L, I) gave (1*S*,2*R*)-**4b** with up to 57% *ee* for CDO M232L while variants CDO M232V, C, S, and G yielded the opposite enantiomer (1*R*,2*S*)-**4b** (Figure 15B). Amongst the mutants that favored the (1*R*,2*S*)-enantiomer of **4b**, CDO M232A still showed the best stereoselectivity with 54% *ee*, followed by CDO M232S and CDO M232G having slightly lower *ee* values (47% and 46% *ee*, respectively, Table 18).

Docking analysis of compound **4** in the active sites of CDO M232L, CDO M232S and CDO M232G, the mutants resulting in the highest stereoselectivities for (1*S*,2*R*)- or (1*R*,2*S*)-**4b** of the examined NDT library variants, indicated different substrate binding modes. While in CDO M232S and CDO M232G, the positioning of the substrate to the catalytic iron center was identical, **4** was bound in a different orientation in CDO M232L, enabling the attack of dioxygen from the opposite face of the substrate (Figure 14). These diverse binding modes caused by the single amino acid exchange presumably turned CDO M232L and CDO M232G/ CDO M232S into enantiocomplementary biocatalysts, yielding opposite absolute configurations for diol **4b**.



**Figure 14:** Docking of indene **4** (green spheres with the C=C double bond shown in dark green) in the active site of CDO M232L (left) as well as of CDO M232S and CDO M232G (right, both structures are aligned). Amino acid residues at position 232 in the CDO variants are shown in red (CDO M232L and CDO M232S) or orange (CDO M232G). Amino acids that constitute the 2-His-1-carboxylate facial triad are shown as grey lines and the catalytic iron as orange spheres. Homology models of the three CDO variants were created with SWISS-MODEL<sup>183</sup> based on the CDO crystal structure (PDB code: 1WQL).<sup>170</sup> Structures of CDO M232S and CDO M232G with **4** docked into the active site were aligned in the PyMOL molecular graphics system, version 1.2r1 (Schrödinger, LLC) and show the same substrate orientation while in CDO M232L, **4** is positioned with its opposite face to the catalytic iron.

In contrast to the stereoselectivities, the effect on mono- versus dihydroxylation for substrate **4** observed within the NDT mutant library seemed, however, not to be clearly correlated with the amino acid side chain size at position 232 (Table 17). Highest ratios of the diol **4b** were obtained with variants CDO M232H (78%), M232F (72%) and M232G (65%) while mutant CDO M232C gave best ratios for the allylic alcohol **4a** (52%).



**Figure 15:** (A) ee Values for the formation of **4a** with CDO, CDO M232A and variants of the CDO NDT mutant library. (B) ee Values for the formation of **4b** with CDO, CDO M232A and variants of the CDO NDT mutant library.

**Table 17:** Product formations and distributions for CDO variants of the NDT mutant library.

CDO variant	Product formation [%] <sup>[a]</sup>		Product distribution [%] <sup>[b]</sup>	
	1 <sup>[c]</sup>	4 <sup>[c]</sup>	1	4
M232H	32 ± 1	3 ± 0.2	14:86 ( <b>1a:1b</b> )	22:78 ( <b>4a:4b</b> )
M232L	29 ± 5	7 ± 2	31:69 ( <b>1a:1b</b> )	42:58 ( <b>4a:4b</b> )
M232F	34 ± 3	5 ± 0.4	28:72 ( <b>1a:1b</b> )	28:72 ( <b>4a:4b</b> )
M232I	46 ± 7	9 ± 1	36:64 ( <b>1a:1b</b> )	50:50 ( <b>4a:4b</b> )
M232V	36 ± 2	4 ± 0.4	54:46 ( <b>1a:1b</b> )	46:54 ( <b>4a:4b</b> )
M232C	42 ± 5	15 ± 1	79:21 ( <b>1a:1b</b> )	52:48 ( <b>4a:4b</b> )
M232S	32 ± 6	19 ± 1	96:4 ( <b>1a:1b</b> )	50:50 ( <b>4a:4b</b> )
M232G	43 ± 6	19 ± 4	95:5 ( <b>1a:1b</b> )	35:65 ( <b>4a:4b</b> )
M232Y	-	-	-	-
M232D	-	-	-	-
M232N	-	-	-	-

<sup>[a]</sup> Product formations were determined by GC-FID analysis after 24 h using standard curves of the respective products. Standard deviations also include variations in the expression level as duplicates from two different expression batches were used for the determination of product formations. <sup>[b]</sup> Product distributions were determined by GC-FID analysis. <sup>[c]</sup> Total product formation (**1a** and **1b** as well as **4a** and **4b**). Traces of 1-indanol were detected (1-2%) resulting from the RO-catalyzed oxidation of 1-indane present in small amounts in **4**. -: not detected. n.d.: not determined.

**Table 18:** Stereoselectivities for CDO variants of the NDT mutant library.

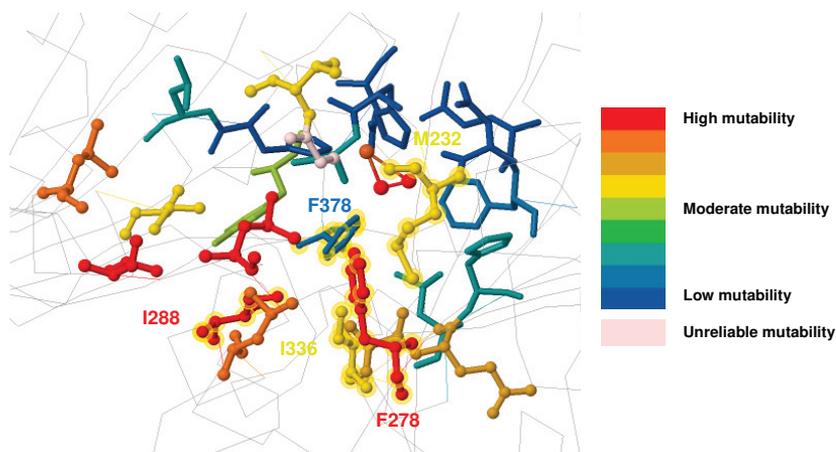
CDO variant	Stereoselectivity ee [%] <sup>[a]</sup>		
	1a	4a	4b
M232H	20 ± 1 ( <i>R</i> - <b>1a</b> )	28 ± 2 ( <i>R</i> - <b>4a</b> )	n.d.
M232L	47 ± 2 ( <i>R</i> - <b>1a</b> )	16 ± 1 ( <i>S</i> - <b>4a</b> )	57 ± 1 (1 <i>S</i> ,2 <i>R</i> - <b>4b</b> )
M232F	36 ± 2 ( <i>R</i> - <b>1a</b> )	36 ± 0.3 ( <i>S</i> - <b>4a</b> )	9 ± 3 (1 <i>S</i> ,2 <i>R</i> - <b>4b</b> )
M232I	61 ± 2 ( <i>R</i> - <b>1a</b> )	52 ± 0.1 ( <i>S</i> - <b>4a</b> )	24 ± 2 (1 <i>S</i> ,2 <i>R</i> - <b>4b</b> )
M232V	65 ± 1 ( <i>R</i> - <b>1a</b> )	81 ± 1 ( <i>S</i> - <b>4a</b> )	12 ± 1 (1 <i>R</i> ,2 <i>S</i> - <b>4b</b> )
M232C	84 ± 3 ( <i>R</i> - <b>1a</b> )	80 ± 1 ( <i>S</i> - <b>4a</b> )	12 ± 1 (1 <i>R</i> ,2 <i>S</i> - <b>4b</b> )
M232S	88 ± 3 ( <i>R</i> - <b>1a</b> )	86 ± 0.1 ( <i>S</i> - <b>4a</b> )	47 ± 1 (1 <i>R</i> ,2 <i>S</i> - <b>4b</b> )
M232G	94 ± 0.3 ( <i>R</i> - <b>1a</b> )	82 ± 0.1 ( <i>S</i> - <b>4a</b> )	46 ± 2 (1 <i>R</i> ,2 <i>S</i> - <b>4b</b> )
M232Y	-	-	-
M232D	-	-	-
M232N	-	-	-

<sup>[a]</sup> Enantiomeric excess values were determined by HPLC analysis using a chiral stationary phase. For chiral HPLC chromatograms see supplementary. Standard deviations were determined from triplicates. -: not detected, n.d.: not determined.

### 3.3.2. Site-saturation mutagenesis at additional positions in CDO

In order to identify further residues that influence the selectivity and activity of ROs, site-saturation mutagenesis at four different positions (F278, I288, I336 and F378) in the substrate binding pocket of CDO was performed. Most of these residues were selected using HotSpot Wizard, a web server for the automatic identification of “hot spots” in order to engineer substrate specificity, activity or enantioselectivity of enzymes.<sup>184</sup> HotSpot Wizard targets amino acids with low conservation levels which are located in the active site or lining the access tunnels. Selecting only functional residues in highly variable positions represents an effective strategy for altering catalytic properties with the reduced risk that the amino acid replacement will compromise the enzymatic function.<sup>184</sup> In contrast to position M232 having an average conservation level, the other residues located within 8 Å of the catalytic iron center (Table 14) were higher conserved amongst ROs<sup>d</sup> and hence not suggested for mutagenesis by the web server. Yet, even though residue F378 was ranked with low mutability, it was nevertheless chosen for site-saturation mutagenesis as substitutions at the corresponding position F352 have been reported to be a major hotspot affecting selectivity in NDO.<sup>124</sup>

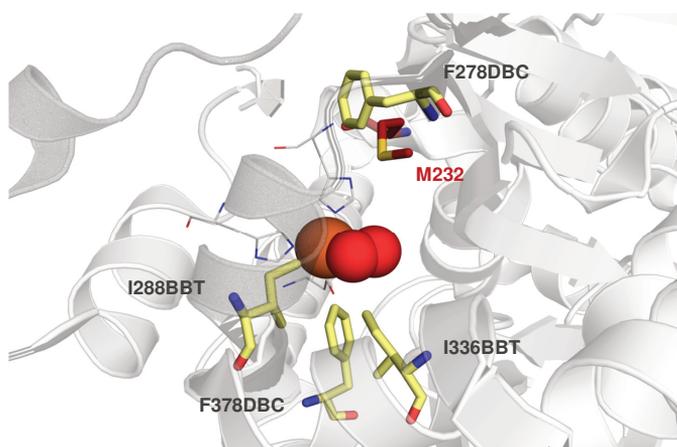
In order to select further residues for mutagenesis, the radius around the non-heme iron was expanded to 12 Å as within this distance, amino acid residues still have an increased probability to result in an altered substrate specificity or stereoselectivity.<sup>123</sup> Using HotSpot Wizard, positions F278, I288 and I336 were identified as residues with high (F278 and I288) and average mutability (I336, Figure 16). In CDO, residues F278 and I288 are part of a loop that ranges from F278 to G290 and runs parallel to a second loop (L248 to N264), opening a channel to the active site. Furthermore, the two loops are also involved in formation of the catalytic pocket, thus representing an interesting target for mutations.<sup>170</sup>



**Figure 16:** Mutability of different functional amino acid residues in CDO as determined by HotSpot Wizard<sup>184</sup> according to their conservation level, ranging from high (red) to low mutability (blue). Amino acids selected for mutagenesis as well as residue M232 are labeled.

<sup>d</sup> The conservation level and therewith the mutability of residues was determined by HotSpot Wizard.<sup>184</sup>

As several alkenes of the selected substrate panel are substituted with sterically demanding methyl groups and lack the planar nature of the natural aromatic substrates, a diminishment of the active site volume was avoided. Therefore, bulky amino acid side chains like tyrosine, tryptophan, lysine or glutamine were excluded from mutagenesis. To cover a broad range of structurally diverse amino acids that were equal in size or smaller than the corresponding residues in the wild type enzyme, BBT and DBC degenerated codons were chosen for library construction. DBC and BBT encode 8 and 9 amino acids (DBC: F, C, I, T, S, V, A, G and BBT: F, S, C, L, P, R, V, A, G), respectively, resulting in only 25 colonies that have to be screened for a library coverage of 95%.<sup>e</sup> Herein, the BBT codon degeneracy was used for positions encoding isoleucine in the wild type enzyme as DBC also contains this amino acid residue. In total, 34 variants were generated by site-saturation mutagenesis within the four mutant libraries that were screened in a first step with a colorimetric solid phase assay for activity.



**Figure 17:** Amino acid residues targeted for site-saturation mutagenesis in CDO (PDB code: 1WQL)<sup>170</sup> using the BBT (F, S, C, L, P, R, V, A and G) and DBC (F, C, I, T, S, V, A and G) codon degeneracy (B = C, G or T; D = A, G or T). Positions F278, I288, I336 and F378 are shown as yellow sticks, position M232 is shown as red sticks. The catalytic non-heme iron center (orange spheres) with bound O<sub>2</sub> (red spheres) is coordinated by the 2-His-1-carboxylate facial triad (shown as grey lines).

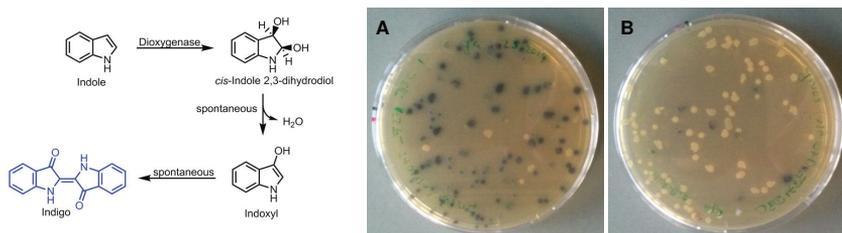
### Colorimetric solid phase screening - Indigo assay

To exclusively select RO mutants for further characterization that have maintained their activity, a solid phase screening with indole was performed. This assay has been applied several times in literature to choose only colonies that turned blue due to the expression of RO mutants still able to oxidize indole and thus form the blue dye indigo (Figure 18).<sup>109,124,185</sup> Also for CDO wild type which was employed as control to test the functionality of the assay, blue color formation could be observed.

<sup>e</sup> The number of colonies that have to be screened for 95% coverage was calculated by CASTER.<sup>223</sup>

Using the solid phase assay, libraries of CDO F278DBC, I288BBT, I336BBT and F378DBC were screened for indole oxidation. Interestingly, variants of the focused mutant libraries based on positions F278, I288 and I336 showed more pronounced indigo formation than the library of CDO F378DBC and about 86-93% of the colonies turned blue during screening. In contrast, the F378DBC mutant library contained only a small fraction of colored colonies (10%, Figure 18 and Table 19) and the blue clones sent for sequencing all represented wild type CDO (DBC also encodes for phenylalanine). Based on observed indigo formation, positions 278, 288 and 336 seemed to be rather robust for mutations as oxidation activity for indole still could be maintained. In contrast, position 378 had a lower tolerance for introduced mutations and variants of this library apparently had lost their ability to convert indole.

Even though the colorimetric solid phase assay allowed the fast screening of a large number of colonies, it was not based on the substrate or product of interest and thus, false negative or positive results might be obtained due to the structural differences between indole and the targeted alkene panel. To examine the suitability of the screening assay for the selection of CDO variants, also white colonies were tested in *in vivo* biotransformations for their activity towards the selected alkenes. Of the examined mutants from the focused libraries F378DBC and I288BBT displaying no indigo formation, neither variant F378S nor mutant I288P showed any activity for olefins **1**, **3**, **4**, **5**, **6** and **7** what corresponds to the lack of color formation observed in the indole assay.



**Figure 18:** Solid phase colorimetric assay based on indigo formation. **(A)** The majority of colonies (> 80%) from mutant libraries based on F278DBC, I288BBT and I336BBT showed blue color formation (DBC library based on position F278 is shown representative for the other mutant libraries). **(B)** Most variants of the mutant library based on position F378 did not show activity towards indole.

**Table 19:** Screening results from the CDO BBT and DBC mutant libraries with the colorimetric solid phase assay based on indigo formation.

Library	Total colonies screened	Number of blue colonies	[%]
CDO F278DBC	108	100	93
CDO I288BBT	100	88	88
CDO I336BBT	154	132	86
CDO F378DBC	90	g <sup>[a]</sup>	10

<sup>[a]</sup> All sequenced blue colonies represented CDO wild type as DBC also encodes for phenylalanine.

## Characterization of CDO mutants

In order to characterize the generated mutant libraries of CDO, the effect of introduced mutations on product formation and selectivity for alkenes **1**, **3**, **4**, **5**, **6** and **7** was tested and the examined variants were compared to CDO wild type. Among the tested mutants, CDO F278A, F278C and I336V displayed altered selectivities and product formations for olefins **1**, **3**, **4** and **5** (Table 20).

**Table 20:** Product formations as well as regio- and stereoselectivities for CDO wt and selected CDO variants.

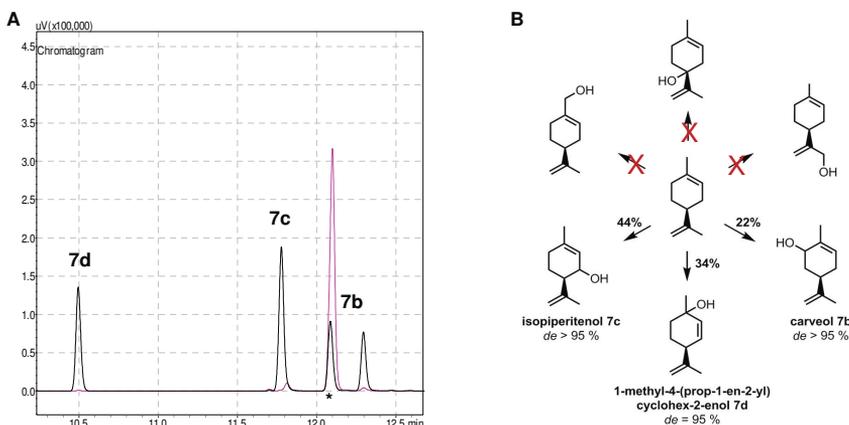
Olefin	CDO wt	CDO F278A	CDO F278C	CDO I336V
Product formation [%] <sup>[a]</sup>				
<b>1</b>	71 ± 6	> 99	57 ± 0.3	45 ± 3
<b>3</b>	36 ± 4	11 ± 0.2	28 ± 6	10 ± 2
<b>4</b> <sup>[e]</sup>	80 ± 7	22 ± 4	18 ± 6	17 ± 0.3
<b>5</b>	29 ± 6	< 1	14 ± 2	13 ± 1
<b>6</b>	-	-	-	-
Stereoselectivity <i>ee</i> [%] <sup>[b]</sup>				
<b>1</b>	43 ± 3 ( <i>R</i> - <b>1a</b> )	84 ± 3 ( <i>R</i> - <b>1a</b> )	70 ± 4 ( <i>R</i> - <b>1a</b> )	n.d.
<b>3</b>	74 ± 0.5 ( <i>R</i> - <b>3a</b> )	56 ± 0.4 ( <i>R</i> - <b>3a</b> )	71 ± 0.4 ( <i>R</i> - <b>3a</b> )	73 ± 1 ( <i>R</i> - <b>3a</b> )
<b>4</b>	38 ± 1 (1 <i>S</i> ,2 <i>R</i> - <b>4b</b> )	12 ± 0.4 (1 <i>S</i> ,2 <i>R</i> - <b>4b</b> )	8 ± 4 (1 <i>S</i> ,2 <i>R</i> - <b>4b</b> )	33 ± 2 (1 <i>S</i> ,2 <i>R</i> - <b>4b</b> )
<b>5</b>	78 (n.d., <b>5b</b> ) > 95 ( <i>R</i> - <b>5c</b> ) 47 (n.d., <b>5d</b> )	n.d.	n.d.	n.d.
<b>6</b>	-	-	-	-
Product distribution [%] <sup>[c]</sup>				
<b>1</b>	0.3:99.7 ( <b>1a:1b</b> )	19:81 ( <b>1a:1b</b> )	18:82 ( <b>1a:1b</b> )	10:90 ( <b>1a:1b</b> )
<b>3</b> <sup>[d]</sup>	80:20 ( <b>3a:3b</b> )	78:22 ( <b>3a:3b</b> )	79:21 ( <b>3a:3b</b> )	78:22 ( <b>3a:3b</b> )
<b>4</b> <sup>[e]</sup>	64:36 ( <b>4a:4b</b> )	50:50 ( <b>4a:4b</b> )	54:46 ( <b>4a:4b</b> )	58:42 ( <b>4a:4b</b> )
<b>5</b>	48:18:18:17 ( <b>5a:5b:5c:5d</b> )	n.d.	59:14:8:19 ( <b>5a:5b:5c:5d</b> )	52:19:10:19 ( <b>5a:5b:5c:5d</b> )
<b>6</b>	-	-	-	-

<sup>[a]</sup> Product formations were determined by GC-FID analysis after 24 h using standard curves of the respective products. <sup>[b]</sup> Enantiomeric excess values were determined by HPLC analysis using a chiral stationary phase. <sup>[c]</sup> Product distributions were determined by GC-FID analysis. <sup>[d]</sup> Cyclohexanone formation resulting from isomerization of **3a** was added to the total amount of **3a**. <sup>[e]</sup> Traces of 1-indanol were detected (1-2%) resulting from the RO-catalyzed oxidation of 1-indane present in small amounts in **4**. To determine standard deviations for the CDO variants, triplicates from one expression batch were used. -: not detected. n.d.: not determined.

With alkene **1**, regioselectivities for products **1a** and **1b** were slightly altered for each of the three mutants with a larger fraction of **1a** being formed compared to the wild type enzyme. Yet, the ratio of **1a:1b** did not exceed 19:81 as obtained for the CDO variant F278A (Table 20), indicating that introduced mutations did not have a pronounced effect on regioselectivity for **1**. In contrast, mutants CDO F278A and CDO F278C showed a strong increase in stereoselectivities for **1a** yielding maximum *ee* values of 84% (*R*)-**1a** for CDO F278A, the variant also giving the highest product

formation (> 99%). For oxidation of alkenes **3–5**, all examined CDO variants showed decreased conversions and/ or stereoselectivities compared to CDO wild type (Table 20). For substrate **3**, stereoselectivities dropped especially after the replacement of wild type amino acid residues with small side chain sizes as observed for CDO F278A (56% *ee* for (*R*)-**3a** compared to 74% *ee*) while for diol **4b**, the decrease of *ee* values was most pronounced with mutants CDO F278A and CDO F278C (Table 20). Yet, absolute configurations of oxidized products did not change compared to the wild type enzyme. Additionally, CDO variants showed slight alterations in regioselectivity for substrates **4** and **5** while ratios of **3a** and **3b** were not influenced (Table 20). In contrast to the CDO variant M232A, none of the mutations enabled the oxidation of terpene **6**.

Whereas variants CDO F278A, CDO 278C and CDO I336V showed the most significant alterations in selectivity for alkenes **1**, **3**, **4** and **5**, a strong shift in regioselectivity for substrate (*S*)-**7** was observed with mutant CDO I288V. Compared to CDO wild type, the main product in the biotransformation of (*R*)-**7** remained unchanged and the allylic alcohol **7b** was formed with good regio- (> 90%, Figure 19A) and stereoselectivity (> 95% *de* for 1*R*,5*S*-**7b**). Yet, for the opposite substrate enantiomer (*S*)-**7**, CDO I288V gave 44% isopiperitenol **7c** and 34% 1-methyl-4-(prop-1-en-2-yl) cyclohex-2-enol **7d** while **7b** was only obtained with 22% (Figure 19). Allylic alcohols **7b** and **7c** were isolated from large scale biotransformations for structure determination by NMR analysis. Product **7d** was identified by comparison with the authentic standard and might result from dehydration of limonene-1,2-diol. Indeed, traces of *cis*-limonene-1,2-diol were obtained during the CDO I288V-catalyzed oxidation of (*S*)-**7**. Whether compound **7d**, however, resulted from dehydration of the diol and whether this reaction might be spontaneous or enzyme-catalyzed was not determined. Even though obtained as product mixture, the stereoselectivities for **7b**, **7c** and **7d** were excellent with *de* values  $\geq$  95%.



**Figure 19:** (A) GC chromatogram of the CDO I288V-catalyzed biotransformation of (*R*)-**7** (pink) and (*S*)-**7** (black). \*: In biotransformations of (*S*)-**7**, (1*R*,5*S*)-**7b** results from oxidation of (*R*)-**7** contaminants that are present in (*S*)-**7** as determined by chiral GC analysis. (B) Product distribution for the oxidation of (*S*)-**7** by CDO I288V. The product ratio only includes the main products **7b**, **7c** and **7d** and several minor potential oxidation products detected in GC-FID analysis that could not be clearly identified due to low amounts were excluded.

### 3.4. Semi-preparative biotransformations with CDO M232A as catalyst

To proof the applicability of ROs for organic synthesis, semi-preparative scale biotransformations of selected substrates were performed with the CDO variant M232A. The mutant was chosen as catalyst due to its improved characteristics compared to the wild type enzymes and other CDO variants. For oxyfunctionalization, compounds **6** and (*R*)-**7** were selected considering the good regio- and stereoselectivities obtained for these terpenes. On the one hand, shake flask experiments with 70 mg substrate were performed using resting cells supplemented with glucose for *in situ* cofactor regeneration. Further up-scaling to 4 L was tested with growing cells in a fed-batch process using a stirred tank bioreactor.

#### 3.4.1. Biotransformations with resting cells in 70 mg scale

Semi-preparative scale biotransformations of substrates **6** and (*R*)-**7** were performed in 70 mg scale with resting cells expressing CDO M232A. For up-scaling of oxyfunctionalizations, a reaction set-up comparable to analytical scale biotransformations was used. Since a sufficient headspace volume proved to be critical for effective substrate oxidation, a ratio of 20:1 between headspace and the reaction mixture was maintained which was also applied for analytical scale oxidations.

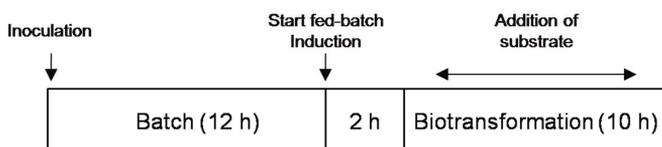
Without further optimization of the reaction set-up, (1*S*,2*S*,3*R*,5*S*)-(+)-**6a** and (1*R*,5*S*)-(+)-**7b** were isolated and purified by flash column chromatography in 33% and 38% yield, respectively (35% and 87% conversion). For **7b**, the decay in isolated yield from 87% product formation to 38% was presumably caused by product loss during extraction and flash column chromatography. On the one hand, in preparative biotransformations, downstream processing of **7b** was found to be difficult as phase separation during extraction was not as easily achieved as in analytical biotransformations. In addition, some fractions of the flash column purification had to be discarded as they contained the allylic alcohol **7a** that could not be completely separated, thereby leading to a further reduced overall yield. Yet, the good regio- and stereoselectivities obtained for products **6a** and **7b** in analytical scale biotransformations were maintained during up-scaling to 70 mg substrate.

#### 3.4.2. Biotransformation of (*R*)-limonene with growing cells in a bioreactor (4 L)

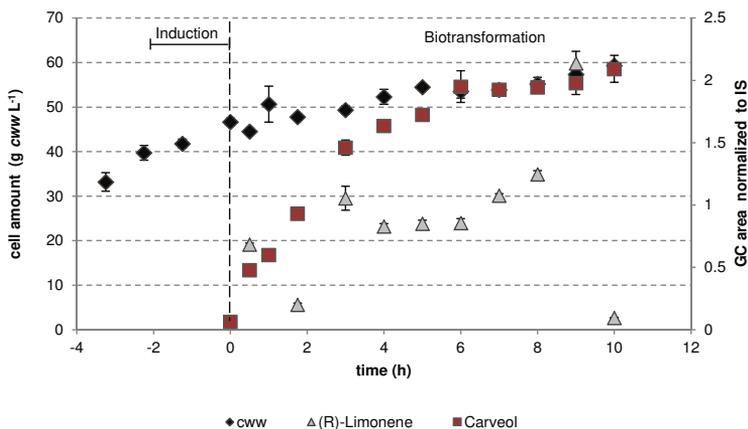
For further up-scaling, hydroxylation of (*R*)-**7** was performed on 4 L scale in a stirred tank bioreactor. To circumvent harvesting of cells and storage of the biocatalyst, the reaction was performed with growing cells having high metabolic activity for cofactor regeneration.<sup>186</sup> The biotransformation immediately proceeded protein expression by direct addition of the substrate to the fermentation broth. Compound (*R*)-**7** was chosen as substrate for scale-up since its oxyfunctionalization by CDO M232A had already been successfully tested in shake flask experiments with 70 mg substrate. Furthermore, its low toxicity in combination with a flash point of 50°C made terpene (*R*)-**7** suitable for experiments performed in a bioreactor. Representing the lowest point at which a compound can vaporize to form an ignitable mixture in air, the flash point of a substrate during biotransformation in a stirred tank reactor should be significantly higher than the reaction temperature to avoid the risk of explosion. As the biotransformation was performed at 30°C, terpene **6** as well as alkenes **1** and **2** also showing good regio- and stereoselectivities in CDO

M232A-catalyzed oxidations, were excluded for scale-up due to their low flash points (**1**: 32°C, **2**: 21°C and **6**: 33°C).<sup>f</sup>

Applying a protocol adapted by Yildirim *et al.*,<sup>154</sup> the fermentation process was divided into three parts (Figure 20). In order to generate cell mass, cells were grown in batch culture over night (~12 h) to obtain about 30 g *cww* L<sup>-1</sup>. Subsequently, fed-batch fermentation was started with a constant feed of glycerol. After induction of protein expression for 2 h, the biotransformation was initiated by addition of (*R*)-**7**. Due to the reported antimicrobial effect of (*R*)-**7** on growing *E. coli* cells,<sup>156</sup> the substrate was added repeatedly during the biotransformation, not exceeding a final concentration of 5 mM. However, due to its high volatility, (*R*)-**7** was stripped off the reactor and had to be added frequently (8 x 5 mM) to obtain a constant product formation.



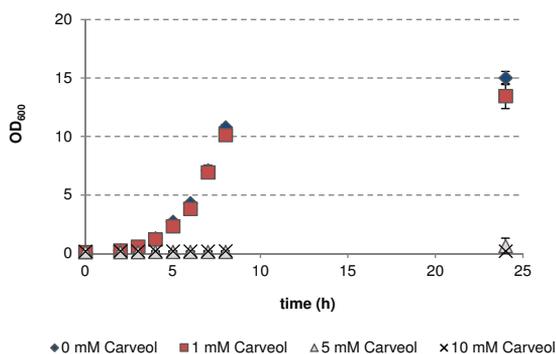
**Figure 20:** Simplified scheme of the biotransformation process, adapted from Bühler *et al.*<sup>186</sup> The batch fermentation over night was subsequently followed by a fed-batch cultivation. After induction of protein expression, the biotransformation was started by the repeated addition of substrate.



**Figure 21:** Biotransformation of (*R*)-limonene (*R*)-**7** with growing *E. coli* JM109 cells expressing CDO M232A in a stirred tank bioreactor on 4 L scale. After batch fermentation to generate cell mass, fed-batch fermentation was started (-3 h) and protein expression was induced for 2 h (-2 h - 0 h, Induction). The dashed line indicates the start of the biotransformation by addition of (*R*)-**7** (0 h). Substrate was added repeatedly (8 x 5 mM) during the course of the biotransformation. IS: internal standard.

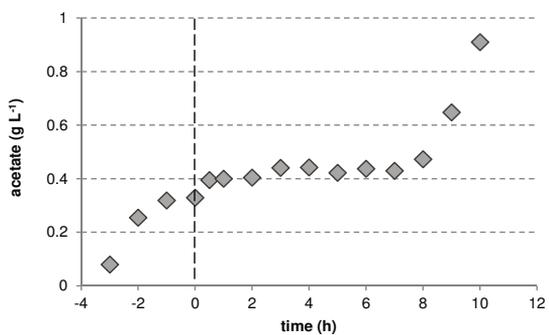
<sup>f</sup> Flash points of the respective compounds were derived from the supplier.

After 10 h biotransformation, 0.4 g L<sup>-1</sup> isolated product yield were obtained, giving (1*R*,5*S*)-(+)-**7b** with an excellent stereoselectivity of 96% *de*. 0.4 g L<sup>-1</sup> **7b** corresponds to a product yield coefficient ( $Y_{P/S}$ ) of only 7% which was partly caused by the volatility of the substrate. Since furthermore, an antimicrobial effect of the product **7b** on growing *E. coli* cells has been reported<sup>187</sup> and the concentration of **7b** did not exceed a certain threshold in the biotransformation (Figure 21), its microbial toxicity was tested for the specific reaction conditions (Figure 22). (1*R*,5*S*)-(+)-**7b** isolated from the fermentation broth was added in varying concentrations (0 mM, 1 mM, 5 mM and 10 mM) to shake flasks inoculated with *E. coli* JM109 cells harboring the plasmid for CDO M232A expression. Starting with an optical density (OD<sub>600</sub>) of 0.05, cell growth was monitored over 24 h at 30°C. The growth rate of cells incubated with 1 mM of **7b** did not differ significantly from that of cells grown without **7b** as a control and both cultures reached an OD<sub>600</sub> of around 15 after 24 h. However, at concentrations of 5 mM and 10 mM **7b**, cells did not show growth within 24 h. Therefore, a limitation of product formation during the biotransformation might be caused by the toxicity of **7b** on growing *E. coli* cells.



**Figure 22:** Microbial toxicity of (1*R*,5*S*)-(+)-carveol **7b** on growing *E. coli* JM109 cells. To determine a toxic effect, cell growth was monitored over 24 h by determining the optical density (OD) of the culture when grown in the presence of different concentrations of (1*R*,5*S*)-(+)-**7b** (0 mM, 1 mM, 5 mM and 10 mM).

Next to product toxicity and substrate volatility, also the formation of acetate by *E. coli* can diminish product yields and detrimental effects on cell growth have been reported for concentrations as low as 0.5 g L<sup>-1</sup>.<sup>188</sup> Therefore, acetate levels were monitored by HPLC analysis during the fed-batch fermentation. Acetate formation increased after the start of the glycerol feed but remained constant at 0.4 g L<sup>-1</sup> when protein expression was induced (Figure 23). Yet, 8 h after the biotransformation was initiated, acetate concentrations started to rise again, yielding acetate levels of 0.9 g L<sup>-1</sup> at the end of the process (10 h).



**Figure 23:** Acetate formation during the fed-batch process with *E. coli* JM109 expressing CDO M232A and a constant glycerol feed over 13 h. After starting the fed-batch process, protein expression was induced for 2 h (-2 h–0 h). At 0 h, the biotransformation was initiated by the addition of (*R*)-limonene (*R*)-7, indicated by the dashed line.

## 4. DISCUSSION

### 4.1. Oxyfunctionalization of alkene substrates by RO wild type enzymes

Based on their activity towards unnatural substrates, three ROs were selected and characterized for the oxyfunctionalization of a set of structurally diverse olefins. Instead of purified enzyme, reactions were performed with whole cells due to the reported instability of isolated ROs<sup>172,173</sup> as well as to provide an efficient *in situ* cofactor regeneration system by the addition of glucose. Examining the selected substrate panel with NDO, BDO and CDO, stereoselectivities and product formations were found to be strongly dependent on the RO as well as the alkene substrate. In order to evaluate the RO-catalyzed oxidation of different alkene types, the olefins comprised four of the six alkenes (mono-, *gem*-di-, *cis*-di- and tri-substituted alkenes). A preference of RO-catalyzed dihydroxylation has been reported by Boyd and coworkers in the order mono-substituted alkene > *cis*-di-substituted alkene > *gem*-di-substituted alkene > tri-substituted alkene > *trans*-di-substituted > tetra-substituted alkene.<sup>74,75</sup> However, with the diverse substrate panel in this study, such general conclusions could be drawn neither with BDO, CDO nor NDO and next to the specific enzyme, also the overall substrate structure as well as the electronic properties of the C=C double bond played an important role for alkene preference. Furthermore, for several of the examined olefin substrates, monohydroxylation in allylic position competed with dihydroxylation of the alkene bond.

As described by Boyd *et al.*, the RO-catalyzed dihydroxylation preferentially occurs at mono-substituted C=C double bonds.<sup>74,75</sup> Yet, even though both compounds contained a mono-substituted alkene bond, styrene **1** was converted with high product formations by all enzymes whereas conversions dropped for vinylcyclohexane **2**, indicating that the substrate preference is not solely determined by the alkene type. On the one hand, the alkene bond of olefin **1** is conjugated in contrast to the unconjugated system of **2**. Consistent with obtained conversions for compound **2**, such unconjugated olefins have already been described to be poor substrates for RO-catalyzed dihydroxylations.<sup>75</sup> Furthermore, alkene **1** has a planar structure due to its aromatic substituent whereas for **2**, the cyclic six-membered ring is supposed to have a chair-like conformation, hence deviating much more from the planar structure of the natural aromatic compounds than **1** (Figure 24).



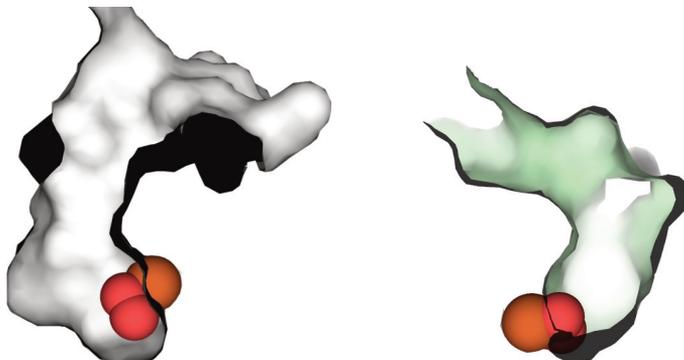
**Figure 24:** 3D-Structures of compounds **1** and **2** generated with Chem3D Pro. In contrast to **1** having a planar aromatic moiety, the cyclohexane ring of **2** adopts an out of plane configuration.

Like compound **2**, the cyclic olefin cyclohexene **3** also possesses an unconjugated yet *cis*-di-substituted C=C double bond. Even though CDO and NDO showed conversions < 40% with **3** as substrate, BDO yielded 59% product formation. Solely relating to activities obtained for the mono-substituted alkene **1**, these results would be in accordance with literature, reporting that mono-substituted olefins are preferred over *cis*-di-substituted alkenes.<sup>74,75</sup> Yet, with BDO, conversion of the mono-substituted alkene **2** was < 1%, indicating that the shape of the active site binding pocket in combination with the spatial structure of the substrate plays a major role for alkene preference. Furthermore, the conjugated *cis*-di-substituted alkene indene **4** was well accepted by the three enzymes and product formations between 80% and > 99% were comparable to conversions obtained for substrate **1** but much higher than for compounds **2** and **3**. Like olefins **2** and **3**, also the five-membered ring of **4** had a twist about the plane compared to aromatic systems. Nevertheless, besides the arene-substituted alkene **1**, compound **4** was the best converted substrate in the panel. Even though allylic monohydroxylation competed with dihydroxylation of the C=C double bond, diol formation with substrate **4** was at least with NDO and BDO still higher than for **2**. These results suggest that a conjugated C=C double bond in combination with an aromatic moiety to give the substrate a more planar character is promoting alkene oxyfunctionalization by ROs.

Due to their structural diversity, it was hard to link the olefins of the substrate panel with respect to a preference for a distinct alkene type. For this purpose, polyene myrcene **5** combining a mono-, *gem*-di- as well as tri-substituted C=C double bond in one molecule was a good model substrate. With NDO, the sequence mono-substituted alkene > *gem*-di-substituted alkene > tri-substituted alkene postulated by Boyd and coworkers could be observed during oxyfunctionalization with a slight preference for the mono-substituted over the *gem*-di-substituted C=C double bond (44:41) whereas dihydroxylation of the tri-substituted alkene bond was clearly disfavored (3%). Yet, the formation of different products was also dependent on the orientation and positioning of **5** in the active site of the three enzymes. With CDO, the allylic alcohol (*Z*)-**5a** was preferred over dihydroxylation and diols **5b-5d** were formed in rather equal ratios even though the terminal mono-substituted and the *gem*-di-substituted C=C double bonds are conjugated and sterically better accessible than the tri-substituted isolated alkene bond giving **5c**. In this enzyme, the shape of the substrate binding pocket might favor orientation of the methyl groups close to the catalytic iron center and thus also show enhanced dihydroxylation of the adjacent tri-substituted alkene bond compared to NDO.

Even though closely related with regard to their overall sequence similarity, such differences in product ratios for BDO, CDO and NDO could be observed for most of the substrates. As already mentioned, the formation of different products indicated diverse binding modes in the active sites of the enzymes due to differently shaped substrate binding pockets (Figure 25). Since the examined olefins lack any oxyfunctionalities, their orientation in the substrate binding pocket is coordinated by hydrophobic interactions and the steric constraints that are specified by the active site topology. Whereas the conversion of their natural arene substrates (benzene, cumene and naphthalene) exclusively yielded the corresponding arene-1,2-dihydrodiols,<sup>189-191</sup> dihydroxylation competed with alcohol formation during oxyfunctionalization of olefins giving varying ratios of mono- and dihydroxylated products. Consistent with previous reports from literature, monohydroxylation reactions were observed to be dependent on both the aliphatic compound as well as the allylic position<sup>73</sup> and monohydroxylated

products were obtained for substrates **3**, **4**, **5** and limonene **7**. Monohydroxylation presumably occurs when the activated allylic carbon center is oriented close to O<sub>2</sub> bound to the catalytic iron and has been suggested to proceed *via* a radical reaction mechanism.<sup>106</sup> While one oxygen atom is inserted during the reaction to yield the alcohol, the second oxygen atom might be reduced to water as observed for monooxygenases like P450 enzymes<sup>192</sup> albeit this has to be elucidated. In contrast to P450 monooxygenases, the oxyfunctionalization of unactivated sp<sup>3</sup> carbon centers was not observed and no example for such a reaction has been described in literature to be catalyzed by ROs.

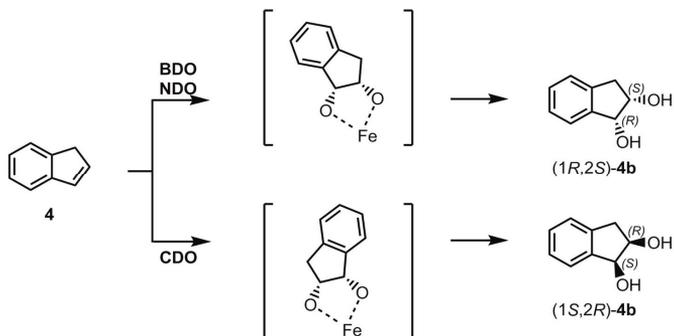


**Figure 25:** Shapes of the substrate binding pockets of CDO (grey, left, PDB code: 1WQL)<sup>170</sup> and NDO (green, right, PDB code: 1O7N)<sup>101</sup>. The catalytic iron and bound oxygen are shown as orange and red spheres.

Also with regard to their stereoselectivities, the three ROs showed interesting differences and opposite enantiomers were obtained for the CDO-catalyzed oxyfunctionalization of **4** compared to NDO and BDO. As mirror-image enzymes for the formation of both product enantiomers are rare in nature, enantiocomplementary biocatalysts performing the same reaction with opposite stereoselectivities are highly interesting. Yet, examples for an enzymatic access to both enantiomers are limited and for ROs, only few reactions have been described, all including NDO and TDO as enantiocomplementary catalysts.<sup>75,175,193,194</sup> Even though the two enzymes share the same protein fold, exchanged locations of their binding sites due to eight amino acids differing in the enzymes' active sites caused the dihydroxylation of 1,2-dihydroxynaphthalene, different aryl alkyl sulfides, indan-2-ol and **4** with opposite stereoselectivities.<sup>75,94</sup> While for the dihydroxylation of **4** the (1*S*,2*R*)-enantiomer of diol **4b** was obtained, the absolute configuration of the allylic alcohol **4a** was dependent on the TDO enzyme. Accordingly, TDO from *P. putida* UV4 gave the opposite enantiomer ((*R*)-**4a**) compared to NDO whereas TDO from *P. putida* 39/D<sup>9</sup> yielded (*S*)-**4a**, too.<sup>90,92,195</sup> Like TDO from the UV4 strain, also CDO showed the opposite enantiopreference for both **4a** and **4b** compared to NDO and BDO. Its reversed enantioselectivity presumably results from different substrate orientations in the active site due to a varying topology of the substrate binding pocket. Thus, in the three enzymes, the active sites are functionally mirror images of one another so that during catalysis of the same reaction, opposite faces of the substrate are exposed to the catalytic iron center resulting in opposite stereoselectivities

<sup>9</sup> *P. putida* 39/D is a mutant of *P. putida* F1 lacking *cis*-dihydrodiol dehydrogenase activity.<sup>65</sup>

(Scheme 18).<sup>94</sup> Interestingly, the  $\alpha$ -subunits of BDO and TDO from *P. putida* F1 containing the active site only differ in 33 amino acid residues<sup>153</sup> even though the two ROs are enantiocomplementary biocatalysts for the formation of **4b**. Similarly, also CDO and BDO share 78% and 63% overall sequence similarity and identity, respectively, while opposite absolute configurations were obtained, indicating that the stereoselectivity of **4** is determined by a small number of amino acid residues in the substrate binding pocket.



**Scheme 18:** Enantiocomplementary dihydroxylation of indene **4** by NDO, BDO and CDO yielding diol **4b** with opposite absolute configurations, modified after Mugford *et al.*<sup>94</sup>

With the exception of **4**, the same absolute configurations were obtained for all other tested substrates in RO-catalyzed biotransformations. In general, with NDO stereoselectivities were considerably higher for most products (**1a**, **2a**, **4a**, **4b**, **5b** and **5d**) compared to BDO and CDO. Only for alcohol **3a**, BDO and CDO showed better *ee* values what might be attributed to the small size of substrate **3**. With the bulky methionine side chain at position 232 and 220, the cyclic olefin **3** could be more tightly bound in one distinct orientation to the iron center in the pockets of both enzymes.

Interestingly, the best stereoselectivities were obtained for linear alkenes which have been described in literature as challenging substrates for RO-catalyzed oxyfunctionalizations,<sup>74,75</sup> yielding good to excellent *ee* values for the dihydroxylation of compounds **2** and **5** (> 95%). In contrast to NDO, CDO gave products **5b** and **5d** with lower ratios and *ee* values of 78% and 47% while stereoselectivities for **5c** were excellent (> 95%). Yet, even the moderate *ee* values for diols **5b** and **5d** were still higher than stereoselectivities reported for shorter acyclic dienes like isoprene, piperylene and 1,3-butadiene with *ee* values not exceeding 45% in whole cell biotransformations using BDO, NDO or TDO.<sup>74</sup>

Regarding the substrate scope, BDO displayed the most limited spectrum for the selected alkene panel. In contrast to NDO or CDO, it did not tolerate several compounds with out of plane configurations or diverse substituents like alkenes **2**, **5**, **7** and 3-chlorostyrene **13** ( $\leq$  5% product formation) being better accepted by the other RO enzymes. Its more restricted substrate specificity might result from the fact that BDO was evolved for the small compound benzene while cumene and naphthalene are more bulky due to the isopropyl moiety or second phenyl ring. Yet, like BDO, also NDO and CDO did not accept (+)- $\alpha$ -pinene **6** and (+)-valencene **8** which represent sterically demanding compounds due to their various methyl substituents and out of plane configurations.

#### 4.2. Influence of different active site residues on the RO-catalyzed alkene oxidation

In contrast to their natural aromatic substrates, the applicability of NDO, CDO and BDO for the oxyfunctionalization of olefins was limited by their poor stereoselectivities and product formations for several of the tested compounds. To overcome these limitations and improve RO enzymes for the conversion of aliphatic substrates, mutations in the binding pockets of the three ROs were introduced. Residue M232 in CDO as well as the corresponding positions in NDO and BDO were selected based on a phylogenetic comparison with other members of the RO family. Furthermore, positions F278, I288, I336 and F378 in CDO were mutated using different codon degeneracies (DBC: F, C, I, T, S, V, A, G and BBT: F, S, C, L, P, R, V, A, G). With the exception of F378 being highly conserved, all targeted positions had an average to high mutability according to HotSpot Wizard, a web server calculating the mutability of a distinct residue from its conservation level.<sup>184</sup> In general, less conserved amino acids are naturally more tolerant to a wider range of substitutions and were thus chosen for mutagenesis.<sup>196</sup> Yet, some positions attributing to enzyme specificity might be conserved to maintain a particular selectivity<sup>196</sup> and since the corresponding residue in NDO (F352) has been reported to have a significant influence on stereo- and regioselectivity for aromatic compounds, position F378 in CDO was selected for mutagenesis.<sup>124,125</sup> However, as predicted by HotSpot Wizard, with variants of the F378DBC mutant library, no activity towards indole was observed in the colorimetric solid phase screening and in line with the assay, clones without color formation that were examined with the substrate panel did not show oxidation activity.

In contrast, most mutants from the libraries based on positions M232, F278, I288 and I336 maintained their activity according to the indole assay (CDO F278DBC, I288BBT and I336BBT) or the substrate screening (CDO M232NDT: F, L, I, V, Y, H, N, D, C, S and G), indicating that these positions are rather robust for the introduction of other amino acids. This tolerance might be attributed to the low conservation level found for the targeted residues, a feature maybe connected to the diverse substrate scope of ROs. Amino acid side chains predicted to interact with the accepted compounds have been reported to be quite variable in dioxygenases whose primary substrates include aromatic hydrocarbons, chlorinated aromatics, amino- and nitroaromatic compounds as well as aromatic acids.<sup>125</sup> These results in combination with previous data suggest that a wide range of amino acid residues can be tolerated near the active site of ROs, having significant effects on the oxyfunctionalization of a variety of compounds<sup>125</sup> as also detected for the examined substrate panel with variants from the generated libraries.

Furthermore, the influence of targeted amino acid residues on selectivities and product formations confirm reports from literature describing that the most effective mutations can be found within 15 Å of a key catalytic atom. All positions that were chosen for mutagenesis are located within a distance of 12 Å of the non-heme iron center as variations of amino acid residues within this distance are more likely to result in an altered substrate specificity or stereoselectivity.<sup>123</sup> In addition, positions F278, I288, I336 and F378 are concentrated at the C-terminal portion of the  $\alpha$ -subunit in CDO which comprises the residues affecting substrate specificity in the toluene/biphenyl RO subfamily<sup>64,197</sup> while M232 is located closer to the active site iron ligands.<sup>182</sup>

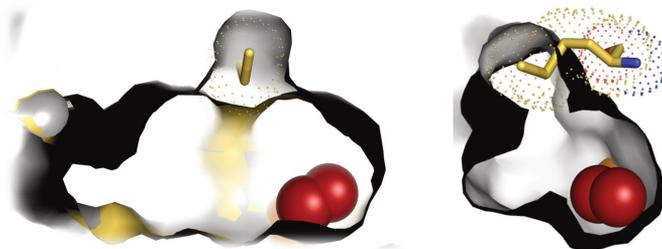
## Mutagenesis of M232 in CDO and its corresponding positions in BDO and NDO

Among the five selected positions in CDO, variation of residue M232 had the most significant effects on selectivity and activity. M232 is one of 14 amino acid side chains constituting the substrate binding pocket and contributes to the more hydrophobic inner surface of CDO compared to NDO.<sup>170</sup> The exchange of M232 in CDO by alanine as found in NDO resulted in a variant with improved selectivities and product formations for several alkene substrates. One of the most significant changes in regio- as well as stereoselectivity was observed for substrate **1**. The single targeted point mutation induced a shift in regioselectivity, yielding alkene-1,2-diol **1a** as main product with enhanced stereoselectivities. By varying the size of the amino acid side chain, regio- and stereoselectivities could be further modulated and variants of the NDT mutant library showed in general higher stereoselectivities and ratios of alkene-1,2-diol **1a** when substituted with smaller residues.

The altered regioselectivity in combination with an increased *ee* value for **1a** was also obtained for BDO M220A. Yet, in contrast to CDO and BDO variants, the reciprocal exchange of alanine by methionine at the corresponding position 206 in NDO did not result in the formation of the alkene-1,2-dihydrodiol. Due to the introduction of the bulky methionine side chain in the substrate binding pocket, however, the activity of NDO mutant A206M dropped significantly. The decrease in conversion might be caused by the diminished active site volume and the replacement of methionine by the smaller isoleucine indeed increased product formations, yet without having an influence on regioselectivity. Similar results have been obtained for the reciprocal exchange of amino acid residues in the active site of closely related P450 monooxygenases to alter the substrate orientation. By phylogenetic comparison, Schalk and Croteau identified a single point mutation that was sufficient to completely switch the regioselectivity for one of the two enzymes.<sup>198</sup> As observed for CDO and NDO, the amino acid substitution did only work in one direction whereas introduction of the larger amino acid residue gave an enzyme inactive for the tested substrate due to compromised binding affinity and a binding orientation that was likely to prevent catalysis.<sup>198</sup>

Besides an increased stereoselectivity for **1a**, enhanced *ee* values with variant CDO M232A were also obtained for products giving better stereoselectivities with NDO than CDO wild type. Stereoselectivities increased to > 95% *ee* for diols **2a** and **5b** and also for the oxyfunctionalized products of **4**, the optical purity was enhanced albeit *ee* values for the diol did not reach stereoselectivities obtained with NDO. Yet, it should be highlighted that the absolute configuration of both **4a** and **4b** was reversed compared to CDO wild type, giving the same absolute configuration as NDO. The switch in stereoselectivity presumably results from different substrate orientations due to an altered active site topology as the amino acid side chain at position 232 forms part of the substrate binding pocket (Figure 26, as no crystal structure of the CDO variant M232A was available, a homology model was created based on CDO wild type). By introducing a single point mutation, enantiocomplementary enzymes for the oxidation of **4** were generated, enabling the attack of dioxygen from opposite faces of the substrate. In the NDT mutant library, opposite stereoselectivities could be further modulated by altering the steric demand of the amino acid side chain at position 232. Smaller side chain sizes (V, C, S, G) increased the *ee* values for (*S*)-**4a** and resulted in an enantiopreference for (1*R*,2*S*)-**4b** while more bulky residues (F, L, I) yielded the (1*S*,2*R*)-enantiomer. An induced switch in stereoselectivity has been already reported for the conversion of aromatic compounds by engineered variants of NDO and

nitrobenzene dioxygenase. Compared to the wild type enzymes, NDO F352V and nitrobenzene dioxygenase variant V207A formed the opposite enantiomers for biphenyl *cis*-3,4-dihydrodiol or different thioanisole derivatives.<sup>124,142</sup> Similar to CDO M232A, in both mutants the original amino acid was substituted with a smaller residue, presumably allowing for opposite substrate orientations to the non-heme iron center in the active sites of the enzymes.



**Figure 26:** Active site model of CDO M232A (grey) in comparison with CDO wild type (yellow). The homology model of the CDO variant M232A was created based on the CDO crystal structure (PDB code: 1WQL)<sup>170</sup> using SWISS-MODEL.<sup>183</sup> The structures of CDO and CDO M232A were aligned in the PyMOL molecular graphics system, version 1.2r1 Schrödinger, LLC. The methionine residue at position 232 is shown as yellow sticks. The catalytic iron center is shown as orange spheres while bound O<sub>2</sub> is shown as red spheres.

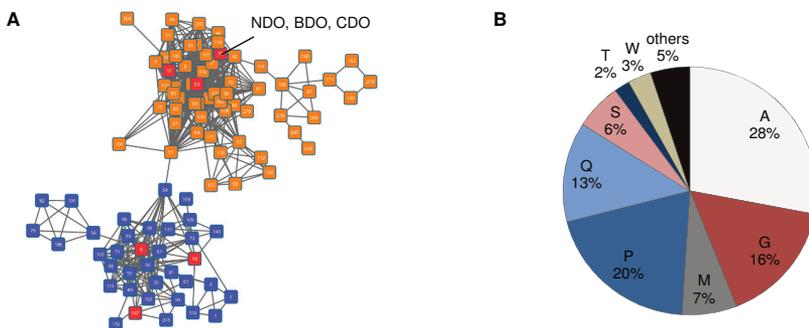
Even though CDO M232A displayed enhanced stereoselectivities for most of the tested olefins, a decrease in *ee* values was obtained for substrates giving lower stereoselectivities with NDO compared to CDO. The drop in *ee* for the allylic alcohol **3a** was presumably caused by the enlarged substrate binding pocket which might allow binding of the small cyclic alkene **3** in more than one defined orientation and was also observed with BDO M220A.

On the other hand, the increase in size of the substrate binding pocket enabled productive binding of the sterically demanding terpene **6** in the active site of CDO M232A. In CDO wild type, accommodation of the substrate might be hindered due to the clash with the bulky methionine residue at position 232. Albeit in NDO, an alanine residue is already present at the corresponding position, other amino acid side chains obviously prevented oxidation of **6** in the enzyme active site. NDO has been described to have a rather flat substrate binding pocket formed by the hydrophobic residues L227, F224 and L253 which can accommodate planar aromatic systems but might impede the binding of substrates with spatial three-dimensional structures.<sup>199</sup> As indicated by docking analysis, also in BDO M220A, productive binding of **6** in the catalytic site might be prevented due to steric constraints caused by amino acid residues lining the entrance tunnel or the substrate binding pocket. In contrast, CDO M232A catalyzed the formation of (1*S*,2*S*,3*R*,5*S*)-**6a** with a high *de* > 95%. The same diastereomer was also obtained in the OsO<sub>4</sub>-catalyzed dihydroxylation of **6** yielding stereoselectivities > 95% *de* even in the absence of a chiral ligand.<sup>163,200</sup> For both the chemical as well as the biocatalyst, the spatial structure of the substrate might prevent an attack from the opposite side of the molecule, thus forming only one of the two possible diastereomers due to steric hindrance.

Besides oxyfunctionalization of terpene **6**, enhanced product formations with CDO M232A were also observed for other substrates. The higher conversions with this CDO variant might be either caused by

higher expression levels, increased stability or enhanced catalytic efficiency. However, due to their cofactor-dependency and multicomponent nature, ROs were only applied in whole cell systems and no  $k_{cat}/K_M$  values were determined in order to compare the catalytic efficiency of wild type enzymes and variants thereof.

Considering the strong effect on selectivity and product formation not only for CDO but also for BDO, position 232 (numbering refers to CDO) was examined in the context of a broader sequence space. In cooperation with Dr. Constantin Vogel from the bioinformatics group of the Institute of Technical Biochemistry, the amino acid distribution at this position as well as the surrounding residues were compared within about 18,000 sequences. According to their global sequence homology, these putative dioxygenases can be grouped into two superfamilies (orange and blue, Figure 27A). The mutated residues in CDO and BDO were found to be located at a variable position (underlined) at the end of a highly conserved motif (N-W-K-x(3)-[ED]-[NQ]-x(3)-[DE]-x-Y-H) which proved to be specific for the superfamily comprising BDO, NDO and CDO (orange, Figure 27A). At this position, alanine was the most frequent amino acid (28%) followed by proline (20%) and glycine (16%, Figure 27B). Its high frequency at position 232 in natural RO variants within the specific superfamily might be one reason why alanine is that well tolerated in CDO.



**Figure 27:** Bioinformatic analysis of the sequence space surrounding position 232 (numbering refers to CDO). **(A)** 18,006 Dioxxygenase sequences were grouped according to their global sequence homology into different homologous families (squares) and two superfamilies (orange and blue). Homologous families displaying more than 40% global sequence similarity are connected by grey lines. Homologous families with available structural information are marked in red. The picture was provided by Dr. Constantin Vogel. **(B)** Amino acid distribution at position 232 within the specific RO superfamily (orange).

Being located at a variable position in a highly conserved motif, residue 232 might also have an influence on selectivity and/ or product formation in other RO enzymes belonging to the same superfamily. Next to CDO and BDO as shown in this study, an effect on selectivity has been demonstrated for both TDO and NDO. The single amino acid exchange M220A at the corresponding position resulted in a TDO variant with activity towards 1,2,4,5-tetrachlorobenzene, a substrate that could not be converted by the wild type enzyme<sup>125</sup> while mutations of A206 in NDO have been reported to show a significant effect on product formations as well as selectivities for several aromatic compounds.<sup>125</sup>

## Mutagenesis of positions F278, I288, I336 and F378 in CDO

In contrast to position M232 that was identified by comparison with another member of the RO family, positions F278, I288, I336 and F378 were selected according to their distance from the catalytic iron (within 12 Å) and their conservation level. To facilitate the identification of active variants within the focused mutant libraries, an initial colorimetric solid phase assay based on indigo formation was applied for screening. This assay has been already described in literature to be suitable for the identification of active clones that did not lose the ability to dihydroxylate indole as a first step towards variants with altered features for the substrate of interest.<sup>109,124,185</sup> Yet, even though indole is structurally similar to compound **4** from the selected substrate panel, it differs greatly from the other olefins that were investigated. Even though colonies that did not show color formation were screened against the alkene panel without detecting activity, a screening assay which employs a different compound than the substrate of interest might nevertheless lead to false negative results. Other assays described for the screening of oxygenases libraries like colorimetric assays based on NADH depletion<sup>139</sup> were however excluded as especially with unnatural substrates, uncoupling has been reported.<sup>201</sup> Thus, such an assay would not only select variants with improved activity towards a distinct substrate but also for increased uncoupling rates. On the other hand, the adrenaline assay for the detection of the diol products did not prove to be suitable as it is sensitive to glucose or glycerol which was used in whole cell biotransformations for *in situ* cofactor regeneration.<sup>202</sup> Despite its disadvantages, the colorimetric solid phase indole assay was found to be applicable for screening of active mutants, being a first indicator which variants had maintained their structural integrity.

As determined with HotSpot Wizard, positions F278, I288, I336 and F378 possess different degrees of conservation. Herein, the conservation level of the residues correlated with the number of colored colonies found in the indole assay. For the DBC library at the highly conserved position F378, the small degree of blue colonies observed during screening (10%) indicated that its tolerance for introduced mutations was rather low. Unlike for CDO, the corresponding position F352 in NDO has been shown to have an effect on both regio- and stereoselectivity for conversion of the aromatic compounds biphenyl and phenanthrene when substituted with other amino acids.<sup>124</sup>

In contrast to position F378, residues F278, I288 and I336 were categorized as average to highly mutable due to their low conservation level. This result was confirmed by the indole assay, indicating active enzyme variants at least for the oxyfunctionalization of indole and > 80% of the screened colonies in these libraries displayed blue color formation. For the selected substrate panel, mutations at position 278, 288 and 336 were furthermore shown to influence selectivities and product formations. In contrast to CDO M232A, only slight changes in regioselectivity could be observed for the oxyfunctionalization of alkene **1** with the examined variants CDO F278A, CDO F278C and CDO I336V. Even though mutants CDO F278A and CDO F278C gave an about 2-fold higher ratio of **1a** than CDO I336V, amounts of the alkene-1,2-diol did not exceed 19%. Yet, CDO variants based on position F278 resulted in a strong increase in stereoselectivity from 43% up to 84% *ee* for (*R*)-**1a**. Accordingly, stereoselectivities for diol **4b** were affected significantly for mutants CDO F278A and CDO F278C, yet, the introduced mutations caused a decrease in *ee* values from 38% to 12% and 8%, respectively, while the absolute configuration of the products was maintained.

The large influence on stereoselectivity of position 278 for the oxyfunctionalization of **1** and **4** might be caused by its close proximity to residue M232 with both residues being located on one side of the catalytic iron center (Figure 17). Furthermore, F278 is located in a loop opening a channel to the active site while forming part of the inner surface of the catalytic pocket.<sup>170</sup> Also residue I288, displaying the most interesting changes in regioselectivity of all examined active variants from the mutant libraries (CDO F278A, CDO F278C, CDO I336V and CDO I288V) forms part of this loop.<sup>170</sup> During oxyfunctionalization of (*S*)-**7**, monohydroxylation still only occurred in allylic position and no oxidation of unactivated C-H bonds was observed. Yet, instead of **7b** as main product, **7c** and **7d** were formed with ratios of 44% and 34%. In contrast to product **7c** obtained by allylic monohydroxylation at C-3 of the substrate, the tertiary alcohol **7d** might be a result of the dehydration of limonene-1,2-diol but no controls have been performed yet to verify this hypothesis. For the opposite substrate enantiomer (*R*)-**7**, the change in regioselectivity was not as pronounced as for (*S*)-**7** and only traces of **7c** could be detected during the CDO I288V-catalyzed biotransformation with **7b** as main product (> 90%). Besides altering the regioselectivity for the oxyfunctionalization of (*S*)-**7**, position I288 has been already shown to influence substrate specificity in natural occurring cumene-type dioxygenase variants from arene-contaminated soil samples.<sup>197</sup> In combination with residue L321, isoleucine at position 288 enabled degradation of toluene and benzene while the substrate scope of variants containing a methionine at both positions was narrowed and toluene was not accepted anymore.<sup>197</sup>

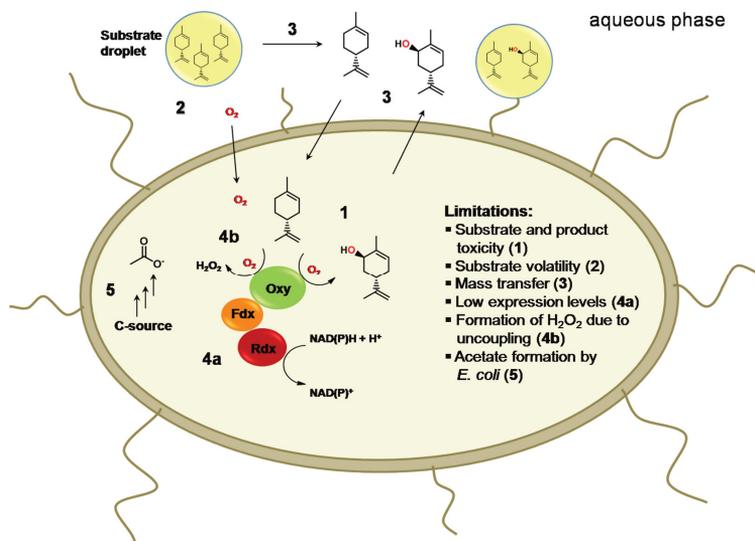
### 4.3. Semi-preparative biotransformations with CDO M232A

As ROs are complex cofactor-dependent multicomponent enzymes, the up-scaling of analytical biotransformations was performed with whole cell biocatalysts for efficient *in-situ* cofactor regeneration and increased oxygenase stability. In two different reaction scales, CDO M232A-catalyzed oxidations were performed with 0.25 g mL<sup>-1</sup> (*cww*) resting cells in shake flask experiments (70 mg of substrates **6** and (*R*)-**7**, respectively) as well as with growing cells in a bioreactor (4 L scale with (*R*)-**7**). Both approaches were compared and advantages as well as disadvantages in the application of resting and growing cells for the oxyfunctionalization of terpene (*R*)-**7** were elucidated.

With 250 g L<sup>-1</sup>, the *cww* in shake flasks experiments was about 5-fold higher than the cell amount in biotransformations of (*R*)-**7** performed with growing cells (between 47 g L<sup>-1</sup> and 59 g L<sup>-1</sup> *cww*). Compared to resting cells giving 27 mg isolated product yield after 24 h (in a total reaction volume of 55 mL), the biotransformation on 4 L scale yielded lower isolated titers of 0.4 g L<sup>-1</sup> **7b** after 10 h. However, considering the significantly lower amount of cells, a higher average specific yield (7.5 mg product g<sup>-1</sup> *cww* versus 1.9 mg product g<sup>-1</sup> *cww*) could be obtained with metabolically active cells.

Despite the higher specific yields, the use of growing cells has disadvantages (Figure 28) and during oxyfunctionalizations, substrate and product toxicity can diminish isolated yields. For *E. coli* JM109 used as expression host, both (*R*)-**7** and **7b** proved to be toxic for growing cells. Due to its reported antimicrobial effect,<sup>156</sup> terpene (*R*)-**7** was added repeatedly in the biotransformation on 4 L scale to maintain concentrations below a toxic level. Whereas detrimental concentrations of the substrate could be avoided, the antimicrobial effect of **7b** that accumulated during the reaction might have reduced the final product titers obtained in the process. As shown in shake flask experiments with growing *E. coli*

JM109 cells, (1*R*,5*S*)-**7b** inhibited cell growth under the specific reaction conditions and after passing a certain threshold in the bioreactor, product formation might be limited by its toxic microbial effect. Besides product toxicity, also substrate volatility severely hampered obtained product yields as during the process, **7** was stripped off the reactor due to air sparging, leading to an overall isolated yield of only 7%. These major constraints of the process, substrate and product toxicity as well as substrate volatility might be overcome by process design. The use of a second liquid phase could not only prevent air stripping of the substrate but also function as a reservoir for antimicrobial compounds like **7** and hence avoid the need of repeated manual addition without inhibiting bacterial cell growth. Therefore, such two-phase systems are frequently employed for the biotransformation of hydrocarbons which often are not only toxic or inhibitory to the biocatalyst but also poorly water soluble.<sup>186</sup> Next to the reduction of substrate loss and toxicity, a second organic phase can also be useful for *in situ* product removal (as well as the use of membrane reactors or solid phase extraction), minimizing the antimicrobial effect of the product to obtain high and stable productivities during cell growth and therewith higher product yields.<sup>154</sup> In such two-phase systems with whole cell RO catalysts, the biocatalyst has been reported to have a greater tolerance to the biotransformation conditions than resting cells when the biotransformation immediately proceeded the fermentation by direct addition of the organic phase to the fermentation broth.<sup>203</sup>



**Figure 28:** Limitations in *in vivo* biotransformations of (*R*)-**7** during fed-batch cultivation of *E. coli* cells in a bioreactor. Figure adapted from Malca.<sup>204</sup>

Due to the above mentioned advantages, oxygenase-based bioprocesses are often carried out in biphasic media.<sup>205</sup> Yet, in such systems, mass transfer of the poorly water soluble hydrocarbon substrate from the organic phase to the biocatalyst is critical and there have been reports in literature when the substrate was trapped in the organic phase and no longer available for the

biotransformation.<sup>206</sup> This problem was observed by Wriessnegger and coworkers for conversion of the sesquiterpene (+)-valencene with *Pichia pastoris*. To eliminate the loss of substrate, *in situ* production of (+)-valencene with valencene synthase was performed by metabolic engineering instead of external substrate addition.<sup>206</sup> As additional alternative, Lynch and coworkers reduced the high loss of substrate during whole cell biotransformations in a bioreactor by replacing air sparging with membrane oxygenation.<sup>207</sup>

Another limitation observed during the process was the formation of acetate which can retard growth and inhibit formation of recombinant protein.<sup>188</sup> Even though glycerol was added as C-source being superior for reduced acetate formation and increased yields of recombinant protein compared to glucose,<sup>188</sup> acetate concentrations of 0.9 g L<sup>-1</sup> were observed at the end of the biotransformation. As detrimental acetate levels have been reported to be as low as 0.5 g L<sup>-1</sup>,<sup>188</sup> a more advanced feeding strategy might be employed to reduce acetate formation and hence enhance product yields. Coupling the feed to dissolved oxygen (DO) levels or a specific growth rate that is maintained below a critical threshold can avoid acetate formation and increase the cell amount. Furthermore, growth rate and acetate formation are strain dependent and also strains with genetic modifications for reduced acetate levels are available.<sup>188</sup>

Compared to RO-catalyzed biotransformations performed in bioreactors on a larger scale, obtained space-time yields for the CDO M232A-mediated oxyfunctionalization of terpene (*R*)-**7** are in the range of other processes with unnatural alkene substrates. Even though reports from literature mainly cover aromatic substrates, several examples are known for the biotransformation of compound **4**.<sup>9,208</sup> For these processes, average space-time yields of 32-50 mg L<sup>-1</sup> h<sup>-1</sup> have been reported employing different biocatalysts which mainly include TDO (Table 21). With around 40 mg L<sup>-1</sup> h<sup>-1</sup>, the space-time yield for oxidation of (*R*)-**7** is comparable to the processes for conversion of **4**. Yet, these biotransformations were performed on a considerably larger scale (15-70 L) and with longer reaction times (Table 21). When compared to up-scaled biotransformations with ROs involving their natural substrates, activities towards aromatics are often considerably higher than for non-physiological compounds (e.g. up to 400 mg L<sup>-1</sup> h<sup>-1</sup> for toluene or naphthalene dependent on the production strain and process conditions).<sup>154</sup>

**Table 21:** RO-catalyzed biotransformations of olefins in a bioreactor adapted from Yildirim *et al.*<sup>154</sup>

RO	Production strain	Substrate	Space-time yield [g L <sup>-1</sup> h <sup>-1</sup> ]	Scale [L]	Time [h]	Reference
CDO M232A	<i>E. coli</i> JM109 (pIP107D_M232A)	( <i>R</i> )- <b>7</b>	0.04	4	10	This study
TDO	<i>P. putida</i>	<b>4</b>	0.044	70 <sup>[a]</sup>	8	9
[b]	<i>Rhodococcus</i> sp. B264-1	<b>4</b>	0.032	23 <sup>[a]</sup>	150	9
TDO	<i>E. coli</i> D160-1	<b>4</b>	0.05	15	24	208
TDO <sup>[a]</sup>	<i>E. coli</i> D160-1	<b>4</b>	0.042	23 <sup>[a]</sup>	24	9

<sup>[a]</sup> total reactor volume, <sup>[b]</sup> not mentioned.

Ranging between 32-50 mg L<sup>-1</sup> h<sup>-1</sup>, average space-time yields obtained in these processes are still below values that would be industrially feasible. Depending on the product, requirements for industrial

oxidation processes include productivities between 10 and 60 g L<sup>-1</sup> d<sup>-1</sup>, high biocatalyst stability, high yields (60-80%), product concentrations of 10-100 g L<sup>-1</sup> and simple product recovery.<sup>186</sup> One main problem for the up-scaling of biotransformations include the reduced specific activities of whole cell biocatalysts which are frequently lower under process conditions in bioreactors than in shake flasks.<sup>154</sup> Due to their multicomponent nature, oxygenases are often instable or inactivated during the bioconversion process.<sup>205</sup> Furthermore, these enzymes have in general low catalytic efficiencies compared with e.g. hydrolases and maximum activities in *in vivo* biotransformations over 10-15 h normally do not exceed 50 U g<sup>-1</sup> *cdw*. Other limiting factors for activity include the transport of substrate and product between the apolar liquid or solid phase and the biocatalyst containing aqueous phase, substrate uptake and product export by whole cell biocatalysts, cofactor regeneration rates, uncoupling as well as oxygenase expression levels in native or heterologous hosts.<sup>205</sup> As also observed for NDO, BDO and CDO, expression levels of ROs are usually low and an increase in protein concentration would be beneficial to obtain significantly higher whole cell oxygenase activities.

#### 4.4. RO-catalyzed alkene oxyfunctionalization *versus* Sharpless AD and Riley oxidation – Benefits and limitations

##### 4.4.1. Sharpless AD

Due to the inherent disadvantages of heavy metal catalysts, the aim of this project was to provide a sustainable biocatalytic alternative to the oxyfunctionalization of alkenes by chemical means. For the AD of olefins, the osmium(VIII)-oxide-catalyzed variant developed by Sharpless and coworkers is still the method of choice due to its broad substrate scope and the good to excellent stereoselectivities that are obtained for five of the six alkene types (mono-, *gem*-di-, *trans*-di-, tri- and tetra-substituted alkenes).<sup>7</sup> Yet, with regard to atom economy, toxicity, side-product formation and catalyst stoichiometry, an enzymatic route would be advantageous.

To be able to compare the RO-mediated dihydroxylation of olefins against Sharpless AD, the selected substrate panel comprised mono- (**1**, **2**, **5**), *gem*-di- (**5**, **7**), *cis*-di- (**3**, **4**) and tri-substituted alkenes (**5**, **6**, **7**). While the examined wild type enzymes still showed major disadvantages compared to the osmium(VIII)-oxide-catalyzed AD, enzyme engineering improved the performance regarding both product formation and selectivity of ROs with *ee* values that were comparable to Sharpless AD for several of the tested olefins (Table 22). Furthermore, next to the selective dihydroxylation of alkene C=C double bonds, also aromatic systems can be regioselectively targeted with RO enzymes and even though the stereoselectivities for product **1b** were not determined, data from literature suggest that such arene-1,2-dihydrodiols are normally formed with excellent *ee* values > 98%.<sup>20,64</sup> In contrast to the AD of alkenes, this reaction is not known to possess a chemical equivalent, further highlighting the potential of ROs for asymmetric synthesis.<sup>51</sup>

From the selected alkene panel, in particular compounds **5** and **6** proved to be interesting substrates for the RO-catalyzed dihydroxylation. Herein, terpene **6** was a challenging substrate not only for Sharpless AD under standard reaction conditions.<sup>163</sup> Also for wild type ROs, the spatial structure of **6** with various methyl groups in combination with the sterically hindered tri-substituted C=C double bond presumably makes this olefin hard to accommodate in a substrate binding pocket evolved for planar

aromatic compounds. Nevertheless, enlarging the size of the active site in CDO M232A enabled productive binding of **6**. With both the biocatalytic as well as the osmium(VIII)-oxide-catalyzed AD, the (1*S*,2*S*,3*R*,5*S*)-(+)-isomer was obtained with excellent stereoselectivities. As with osmium(VIII)-oxide, no chiral ligands were required to induce the high stereoselectivities,<sup>163,200</sup> the steric hindrance of the substrate might favor attack from only one side, promoting the high *de* values obtained with chemical as well as biocatalytic means. For polyene **5**, the regioselectivity of Sharpless AD is determined both by electronic and steric effects. Herein, electronic factors greatly influence the regioselectivity and the osmylation of unsymmetrical polyenes preferentially occurs at the most electron-rich double bond.<sup>7,209</sup> Due to the +I-effect, the tri-substituted C=C double bond of **5** was attacked and diol **5b** was the only product obtained. In contrast, in RO-catalyzed biotransformations, all alkene bonds of the substrate could be targeted yielding excellent stereoselectivities > 95% with the different enzymes (Table 22). Yet, regioselectivities were rather poor and will have to be improved as product mixtures were obtained. The low regioselectivities, partially caused by the RO-catalyzed formation of allylic alcohols, was one of the major disadvantages when comparing the biocatalytic and Sharpless AD. Yet, as shown for the dihydroxylation of **1**, regioselectivities can be improved by enzyme engineering as they are dependent on the positioning and orientation of the substrate in the active site of the enzyme. In contrast, regioselectivities for the Sharpless AD are not only dependent on steric effects but also greatly rely on the electronic properties of the substrate. Due to its electrophilic nature, osmium(VIII)-oxide will preferentially attack the most electron rich double bond when applied under heterogeneous conditions with potassium ferricyanide as oxidant.<sup>7</sup>

**Table 22:** Stereoselectivities obtained by Sharpless AD employing the most suitable chiral ligands compared to best stereoselectivities determined with examined ROs.

Product	Stereoselectivities <i>ee</i> or <i>de</i> [%]	
	RO-catalyzed AD	Sharpless AD
<b>1a</b>	95 ( <i>R</i> )	99 ( <i>R</i> ) <sup>176</sup>
<b>2a</b>	> 98 ( <i>R</i> )	96 ( <i>R</i> ) <sup>176,177</sup>
<b>4b</b>	76 (1 <i>R</i> ,2 <i>S</i> )	63 (1 <i>R</i> ,2 <i>S</i> ) <sup>177</sup>
	57 (1 <i>S</i> ,2 <i>R</i> )	34 (1 <i>S</i> ,2 <i>R</i> ) <sup>23</sup>
<b>5b</b> <sup>[a]</sup>	> 95 (n.d.)	-
<b>5c</b>	> 95 ( <i>R</i> )	> 95 <sup>[b]</sup>
<b>5d</b> <sup>[a]</sup>	> 95 (n.d.)	-
<b>6a</b>	> 95 (1 <i>S</i> ,2 <i>S</i> ,3 <i>R</i> ,5 <i>S</i> )	> 95 (1 <i>S</i> ,2 <i>S</i> ,3 <i>R</i> ,5 <i>S</i> ) <sup>[b]</sup>

<sup>[a]</sup>KMnO<sub>4</sub>-mediated dihydroxylation gives the racemic mixture and thus, absolute configurations could not be determined. <sup>[b]</sup>this study.  
 -: the product was not formed in the OsO<sub>4</sub>-catalyzed AD under standard conditions.

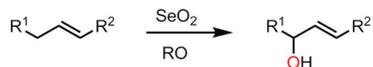
Whereas *cis*-di-substituted alkenes can be oxidized with low to moderate stereoselectivities by Sharpless AD even when employing the best suitable chiral ligands,<sup>24,177</sup> examples have been reported in literature for the RO-catalyzed dihydroxylation of these alkenes with excellent stereoselectivities (e.g. methylindene or cyclohexadiene).<sup>75</sup> However, the *cis*-di-substituted alkene **4**

was a challenging substrate for RO-catalyzed biotransformations and highest stereoselectivities for **4b** were obtained with NDO giving 76% *ee*. Despite these moderate *ee* values, the RO-catalyzed oxidation of **4** proved to be interesting as both enantiomers could be targeted, normally representing a weakness of biocatalysis in asymmetric synthesis.<sup>52</sup> Like AD-mix  $\alpha$  or  $\beta$  which can be employed depending on the desired enantiomer, ROs and generated variants thereof gave comparable selectivities to Sharpless AD for both enantiomers (Table 22).

Considering that Sharpless AD has been continuously improved for decades with regard to its ligands, oxidants and process conditions,<sup>4,7,210</sup> with further engineering it might also be possible to turn ROs into more efficient catalyst for the AD of olefins. Even though in this project, only a limited number of seven alkenes was examined, the selected compounds included structurally highly diverse olefins with different alkene types. For this panel, it was demonstrated that the introduction of a single point mutation proved to be sufficient to transform wild type ROs into efficient catalysts with enhanced stereoselectivities and product formations. Furthermore, the feasibility of engineered RO variants for semi-preparative biotransformations could be shown. Especially the up-scaling of the Sharpless AD is difficult regarding the formation of undesired waste products. Even though both AD-mixes contain osmium(VIII)-oxide in its less volatile form as potassium salt, potassium ferricyanide employed as oxidant for catalyst regeneration is also a highly toxic waste product in the process.<sup>11</sup>

#### 4.4.2. Riley oxidation

Besides the AD of C=C double bonds, the oxidation of allylic carbon centers is catalyzed during the RO-mediated oxyfunctionalization of alkenes. This reaction is of interest as the allylic hydroxylation of C-H bonds can chemically still only be reliably achieved by stoichiometric amounts of SeO<sub>2</sub>.<sup>211</sup> Next to its toxicity, SeO<sub>2</sub>-mediated reactions show poor regioselectivity for substrates with multiple allylic positions displaying similar electronic properties. Therefore, the regiospecific introduction of hydroxyl groups into **7** by chemical means has proven difficult due to the similarity of the allylic methylene (C-3 and C-6) and methyl groups (C-7 and C-10), yielding a product mixture of limonene-4-ol, *trans*- and *cis*-carveol, perillyl alcohol and limonene-10-ol.<sup>212</sup> In contrast, the RO-catalyzed allylic hydroxylation of (*R*)-**7** occurred with good regioselectivity and excellent stereoselectivity at C-6 of the substrate.



**Scheme 19:** Allylic hydroxylation of aliphatic substrates either catalyzed by SeO<sub>2</sub> (Riley oxidation) or by ROs.

Employing less challenging substrates than terpene **7**, the SeO<sub>2</sub>-mediated allylic oxidation showed improved regioselectivities and oxidation of olefin **5** gave exclusively the (*E*)-isomer of alcohol **5a** as generally described for tri-substituted alkenes.<sup>213</sup> In contrast, the biocatalytic oxidation yielded the (*Z*)-isomer, a compound naturally present in thyme oil.<sup>161</sup>

However, due to competing diol formation for several of the examined substrates, ROs still display disadvantages concerning the allylic oxidation of alkenes. While most of the research concerning these enzymes has focused on the AD of substrates, the exact mechanism as well as factors

controlling monohydroxylation *versus* dihydroxylation still have to be elucidated. Except for the oxyfunctionalization of (*R*)-**7** yielding **7b** with good regioselectivity, the allylic oxidation by ROs needs to be further improved to avoid the formation of product mixtures. Yet, with further enzyme engineering to control the regioselectivity of the reaction, ROs might have the potential to also become a sustainable alternative to Riley oxidation.

## 5. CONCLUSIONS AND OUTLOOK

Despite their broad substrate scope, most of the research concerning RO-catalyzed oxyfunctionalizations has focused on the dihydroxylation of aromatic compounds and limited data is available for the oxidation of unnatural substrates by this class of enzymes. Thus, in the present work, the applicability of ROs for the oxidation of a broad range of structurally diverse olefins was demonstrated. Their large hydrophobic substrate binding pocket enables these enzymes to convert a variety of alkenes, lacking the planar structure of the aromatic compounds that are naturally oxidized. This active site feature might account for the versatility of ROs with respect to substrate specificity but also catalyzed reactions. Different substrate positioning in the active site with either the C=C double bond or the allylic/ benzylic carbon of the substrate oriented close to the non-heme iron center presumably results in the formation of diverse reaction products.

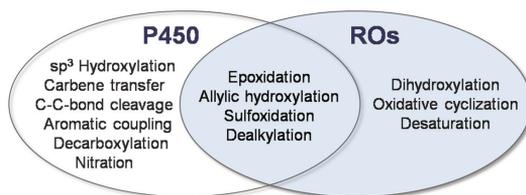
The low stereoselectivities and conversion rates displayed by the wild type enzymes could be overcome by the introduction of a single targeted point mutation that was identified based on sequence and structural comparisons with a member of the RO family. The generated BDO and CDO variants displayed remarkable changes in regio- and stereoselectivities for different aliphatic compounds. Especially the CDO variant M232A resulted in excellent *ee* values  $\geq 95\%$  for linear alkenes, which have been reported to be challenging substrates for RO-catalyzed oxyfunctionalizations.<sup>74,75</sup> Site-saturation mutagenesis of position 232 in CDO using the NDT codon degeneracy revealed a correlation between the steric demand of the introduced amino acid residue and its influence on regio- and/ or stereoselectivity for the tested olefin substrates. For indene biooxidation, enantiocomplementary enzymes were generated by the single point mutation, overcoming the weakness of biocatalysts that lack the presence of mirror-image enzymes for the formation of either enantiomer of a product.<sup>94</sup> These data support the importance of position M232 for the engineering of ROs, a fact that is further highlighted by its location at a variable position (underlined) in the conserved motif N-W-K-x(3)-[ED]-[NQ]-x(3)-[DE]-x-Y-H which is specific for this RO superfamily.

Biocatalysts that accept a wide range of substrates to regioselectively form optically pure products are of particular interest for synthetic organic chemistry.<sup>144</sup> Yet, for compounds like cyclohexene, myrcene or indene, the regio- as well as stereoselectivities obtained with RO enzymes are still a subject for further improvements and additional positions for mutagenesis should be identified to increase the selectivity for the oxidation of alkenes. One starting point towards more selective ROs might be the mutation of the remaining variable residues found in the conserved motif. Since these positions differ within the specific RO superfamily, they might contribute to diverse selectivities or activities. Using an alanine scanning approach, their influence on enzyme properties can be determined. In addition, the combinatorial active-site saturation test (CAST)<sup>214</sup> could be performed to examine additive or synergistic effects, neighboring positions of M232 might have on enzyme properties in order to generate variants with increased regioselectivity for the above mentioned substrates. Modification of the active site topology to allow the binding of these compounds in only one defined orientation would give the desired products with high regio- and stereospecificities. Furthermore, mono- and

dihydroxylation of substrates might be controlled by creating variants that only catalyze one reaction type, yielding either the allylic alcohol or the diol.

Next to enzyme engineering, mono- *versus* dihydroxylation might also be guided by substrate design. Interestingly, for the conjugated olefin 1,3-cyclohexadiene, no competing allylic monohydroxylation has been described for biotransformations with NDO.<sup>75</sup> This result is in contrast to the NDO-mediated oxidation of cyclohexene that was tested in this work and yields both the diol as well as the allylic alcohol, indicating that the type of catalyzed reaction might be influenced by the substrate structure.

Due to their versatile reaction spectrum and broad substrate scope, ROs have often been described as the non-heme iron analogues to cytochrome P450 monooxygenases.<sup>215</sup> Even though several of the different reaction types including sulfoxidations, dealkylations, epoxidations and allylic hydroxylations are catalyzed by both enzyme classes, other reactions performed by P450 monooxygenases cannot be targeted with ROs to date (Figure 29).<sup>93,216,217</sup> By substrate screening with different ROs in combination with enzyme engineering or directed evolution, RO enzymes catalyzing novel types of reactions like the oxidation of unactivated C-H bonds might be generated.



**Figure 29:** Reaction types catalyzed by P450 monooxygenases and ROs.<sup>93,93,216,217</sup> While the reaction spectrum of both enzyme classes partly overlaps, other reaction types are only catalyzed by either P450 monooxygenases or ROs.

Besides their broad substrate and reaction scope, another feature shared by ROs and P450 monooxygenases is the uncoupling of NAD(P)H consumption and substrate oxidation in the presence of compounds lacking a productive binding mode in the active site.<sup>192,201</sup> Next to H<sub>2</sub>O<sub>2</sub> formation which can inactivate the enzyme, uncoupling also leads to a loss of reducing equivalents from the host cell and an increased oxygen demand as well as a lowered specific activity with regard to product formation.<sup>205</sup> For the oxyfunctionalization of the alkene substrate panel examined in this study, the extent of H<sub>2</sub>O<sub>2</sub> production was not determined. However, as uncoupling has been described to constitute ~40-50% of the total O<sub>2</sub> consumption for benzene oxidation with NDO,<sup>201</sup> it might also occur for the reactions performed within this study. Whereas for ROs no data is available from literature, the reduction of uncoupling due to inappropriate positioning of the substrate in the enzyme active site by protein engineering or directed evolution has been reported for P450 monooxygenases.<sup>205</sup> This approach might also be applicable to ROs, diminishing potential H<sub>2</sub>O<sub>2</sub> formation during the biotransformation of olefin substrates and hence increase product formation.

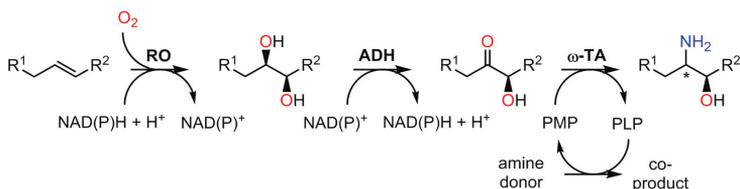
Another strategy to enhance product formation having been described for P450 monooxygenases is the generation of functional fusion proteins between the redox partners and the oxygenase. Linking the single components for more efficient electron transport or higher expression levels due to a better

coordination of transcription and translation has been successfully applied, resulting in an increased substrate oxidation.<sup>218</sup> Also for CDO, a covalent linkage of the reductase and ferredoxin on the gene level proved to be successful and enzyme activity could be reconstituted with the fused redox partners (data not shown). Thus, linkage of the redox construct with the terminal oxygenase component might be performed and compared with the physiological three component system.

In general, the choice of the redox partners (physiological or artificial) and their fusion to the oxygenase component has been assumed not to affect the type and selectivity of reactions catalyzed by P450 monooxygenases. Yet, the group of Li recently reported different reaction types dependent if the P450 enzyme was linked with its surrogate redox partner or present as free component.<sup>216</sup> This example highlights the potential role of redox partner protein-protein-interactions in modulating the catalytic activity of P450 monooxygenases<sup>216</sup> and might also be applicable to other oxygenase multicomponent enzymes like ROs.

Next to biotransformations in analytical scale, the preparative oxidation of (+)- $\alpha$ -pinene and (*R*)-limonene with isolated product yields in mg and g scale was successfully performed. These results indicate that ROs can provide a biocatalytic access to both the asymmetric *cis*-dihydroxylation and the regio- and stereoselective allylic monohydroxylation of various alkenes, targeting also products that are difficult to access by Sharpless AD or Riley oxidation. Several of the biotransformation products generated in this study are interesting building blocks for organic synthesis or might be applied in the flavor and fragrance industry, including carveol, the biooxidation products of myrcene as well as pinanediol.<sup>161,163,178</sup>

Besides *cis*-diols and allylic alcohols, the formation of additional building blocks might be targeted by the integration of ROs into enzyme cascades (Scheme 20). Starting from alkenes, the combination of ROs with ADHs to selectively obtain the corresponding hydroxy ketones from *cis*-diols would be one possibility. These oxidations might be further coupled with  $\omega$ -transaminases ( $\omega$ -TAs) to yield optically pure amino alcohols in a process similar to the work described by Kroutil and coworkers.<sup>219</sup>



**Scheme 20:** Enzymatic cascade involving ROs, ADHs and  $\omega$ -TAs for the generation of amino alcohols starting from alkenes. PLP: pyridoxal-5'-phosphate, PMP: pyridoxamine-5'-phosphate.

Even though still not always competitive with respect to selectivity and product formation, with further optimization, engineered ROs can provide an enzymatic approach to valuable building blocks for the chemical and pharmaceutical industry. Therefore, these enzymes offer a sustainable alternative to existing chemical routes as targeted by CHEM21.

## 6. REFERENCES

1. Chow, T. W.-S., Liu, Y. & Che, C.-M. Practical manganese-catalysed highly enantioselective *cis*-dihydroxylation of electron-deficient alkenes and detection of a *cis*-dioxomanganese(V) intermediate by high resolution ESI-MS analysis. *Chem. Commun.* **47**, 11204–11206 (2011).
2. Punniyamurthy, T., Velusamy, S. & Iqbal, J. Recent advances in transition metal catalyzed oxidation of organic substrates with molecular oxygen. *Chem. Rev.* **105**, 2329–2363 (2005).
3. Chang, D., Zhang, J., Witholt, B. & Li, Z. Chemical and Enzymatic Synthetic Methods for Asymmetric Oxidation of the C–C Double Bond. *Biocatal. Biotransform.* **22**, 113–131 (2004).
4. Zaitsev, A. B. & Adolffson, H. Recent developments in asymmetric dihydroxylations. *Synthesis* 1725–1756 (2006).
5. Bataille, C. J. R. & Donohoe, T. J. Osmium-free direct *syn*-dihydroxylation of alkenes. *Chem. Soc. Rev.* **40**, 114–128 (2011).
6. Bhowmick, K. C. & Joshi, N. N. Syntheses and applications of C<sub>2</sub>-symmetric chiral diols. *Tetrahedron: Asymmetry* **17**, 1901–1929 (2006).
7. Kolb, H. C., VanNieuwenhze, M. S. & Sharpless, K. B. Catalytic Asymmetric Dihydroxylation. *Chem. Rev.* **94**, 2483–2547 (1994).
8. Choudary, B. M., Chowdari, N. S., Madhi, S. & Kantam, M. L. A Trifunctional Catalyst for One-Pot Synthesis of Chiral Diols via Heck Coupling-N-Oxidation-Asymmetric Dihydroxylation: Application for the Synthesis of Diltiazem and Taxol Side Chain. *J. Org. Chem.* **68**, 1736–1746 (2003).
9. Buckland, B. C., Drew, S. W., Connors, N. C., Chartrain, M. M., Lee, C., Salmon, P. M., Gbewonyo, K., Zhou, W., Gailliot, P., Singhvi, R., Olewinski, R. C., Sun, W.-J., Reddy, J., Zhang, J., Jackey, B. A., Taylor, C., Goklen, K. E., Junker, B. & Greasham, R. L. Microbial Conversion of Indene to Indandiol: A Key Intermediate in the Synthesis of CRIXIVAN. *Metab. Eng.* **1**, 63–74 (1999).
10. Döbler, C., Mehlretter, G. & Beller, M. Atom-Efficient Oxidation of Alkenes with Molecular Oxygen: Synthesis of Diols. *Angew. Chem. Int. Ed.* **38**, 3026–3028 (1999).
11. Tse, M. K., Schröder, K. & Beller, M. in *Modern Oxidation Methods* 1–36 (Wiley-VCH Verlag GmbH & Co. KGaA, 2010).
12. Husain, S. M., Stillger, T., Dünkelmann, P., Lödige, M., Walter, L., Breiting, E., Pohl, M., Bürchner, M., Krossing, I., Müller, M., Romano, D. & Molinari, F. Stereoselective Reduction of 2-Hydroxy Ketones towards *syn*- and *anti*-1,2-Diols. *Adv. Synth. Catal.* **353**, 2359–2362 (2011).
13. Notz, W. & List, B. Catalytic Asymmetric Synthesis of *anti*-1,2-Diols. *J. Am. Chem. Soc.* **122**, 7386–7387 (2000).
14. Monaco, M. R., Prévost, S. & List, B. Organocatalytic Asymmetric Hydrolysis of Epoxides. *Angew. Chem. Int. Ed.* **53**, 8142–8145 (2014).
15. Clapés, P., Fessner, W.-D., Sprenger, G. A. & Samland, A. K. Recent progress in stereoselective synthesis with aldolases. *Curr. Opin. Chem. Biol.* **14**, 154–167 (2010).
16. Kurina-Sanz, M., Bisogno, F. R., Lavandera, I., Orden, A. A. & Gotor, V. Promiscuous Substrate Binding Explains the Enzymatic Stereo- and Regiocontrolled Synthesis of Enantiopure Hydroxy Ketones and Diols. *Adv. Synth. Catal.* **351**, 1842–1848 (2009).
17. Shing, T. K. M., Tam, E. K. W., Tai, V. W.-F., Chung, I. H. F. & Jiang, Q. Ruthenium-Catalyzed *cis*-Dihydroxylation of Alkenes: Scope and Limitations. *Chem. Eur. J.* **2**, 50–57 (1996).
18. Jakobinnert, A. & Rother, D. A two-step biocatalytic cascade in micro-aqueous medium: using whole cells to obtain high concentrations of a vicinal diol. *Green Chem.* **16**, 3472–3482 (2014).
19. Fessner, W.-D. in *Enzyme Catalysis in Organic Synthesis* 855–917 (Wiley-VCH Verlag GmbH & Co. KGaA, 2012).

20. R. Boyd, D. & N. Sheldrake, G. The dioxygenase-catalysed formation of vicinal *cis*-diols. *Nat. Prod. Rep.* **15**, 309 (1998).
21. Wackett, L. P. Mechanism and applications of Rieske non-heme iron dioxygenases. *Enzyme Microb. Technol.* **31**, 577–587 (2002).
22. Strassner, T. in *Computational Modeling of Homogeneous Catalysis* **25**, 253–268 (Springer US, 2002).
23. Spivey, A. C., Hanson, R., Scora, N. & Thorpe, S. J. Sharpless Asymmetric Dihydroxylation: Effect of Alkene Structure on Rates and Selectivity An Undergraduate Organic Laboratory Group Experiment. *J. Chem. Educ.* **76**, 655–659 (1999).
24. Wang, L. & Sharpless, K. B. Catalytic Asymmetric Dihydroxylation of *cis*-Disubstituted Olefins. *J. Am. Chem. Soc.* **114**, 7568–7570 (1992).
25. Gawley, R. E. & Aubé, J. in *Principles of Asymmetric Synthesis* 491–544 (Elsevier, 2012).
26. Deubel, D. V & Frenking, G. [3+2] versus [2+2] addition of metal oxides across C=C bonds. Reconciliation of experiment and theory. *Acc. Chem. Res.* **36**, 645–651 (2003).
27. Hoi-Lun, K., Sorato, C., Ogino, Y., Hou, C. & Barry Sharpless, K. Preclusion of the 'second cycle' in the osmium-catalyzed asymmetric dihydroxylation of olefins leads to a superior process. *Tetrahedron Lett.* **31**, 2999–3002 (1990).
28. Wai, J. S. M., Marko, I., Svendsen, J. S., Finn, M. G., Jacobsen, E. N. & Sharpless, K. B. A Mechanistic Insight Leads to a Greatly Improved Osmium-Catalyzed Asymmetric Dihydroxylation Process. *J. Am. Chem. Soc.* **111**, 1123–1125 (1989).
29. Junttila, M. H. & Hormi, O. O. E. Methanesulfonamide: a Cosolvent and a General Acid Catalyst in Sharpless Asymmetric Dihydroxylations. *J. Org. Chem.* **74**, 3038–3047 (2009).
30. Sharpless, K. B., Amberg, W., Bennani, Y. L., Crispino, G. A., Hartung, J., Jeong, K. S., Kwong, H. L., Morikawa, K. & Wang, Z. M. The Osmium-Catalyzed Asymmetric Dihydroxylation: A New Ligand Class and a Process Improvement. *J. Org. Chem.* **57**, 2768–2771 (1992).
31. Xu, D., Crispino, G. A. & Sharpless, K. B. Selective Asymmetric Dihydroxylation of Dienes. *J. Am. Chem. Soc.* **114**, 7570–7571 (1992).
32. Sharpless, K. B. Searching for New Reactivity (Nobel Lecture). *Angew. Chem. Int. Ed.* **41**, 2024–2032 (2002).
33. Plietker, B., Niggemann, M. & Pollrich, A. The acid accelerated ruthenium-catalysed dihydroxylation. Scope and limitations. *Org. Biomol. Chem.* **2**, 1116–1124 (2004).
34. Dash, S., Patel, S. & Mishra, B. K. Oxidation by permanganate: synthetic and mechanistic aspects. *Tetrahedron* **65**, 707–739 (2009).
35. Plietker, B. & Niggemann, M. An Improved Protocol for the RuO<sub>4</sub>-Catalyzed Dihydroxylation of Olefins. *Org. Lett.* **5**, 3353–3356 (2003).
36. Plietker, B. & Niggemann, M. RuCl<sub>3</sub>/CeCl<sub>3</sub>/NaIO<sub>4</sub>: A New Bimetallic Oxidation System for the Mild and Efficient Dihydroxylation of Unreactive Olefins. *J. Org. Chem.* **70**, 2402–2405 (2005).
37. Singh, N. & Lee, D. G. Permanganate: A Green and Versatile Industrial Oxidant. *Org. Process Res. Dev.* **5**, 599–603 (2001).
38. De Vos, D. E., de Wildeman, S., Sels, B. F., Grobet, P. J. & Jacobs, P. A. Selective Alkene Oxidation with H<sub>2</sub>O<sub>2</sub> and a Heterogenized Mn Catalyst: Epoxidation and a New Entry to Vicinal *cis*-Diols. *Angew. Chem. Int. Ed.* **38**, 980–983 (1999).
39. Brinksma, J., Schmieder, L., Van Vliet, G., Boaron, R., Hage, R., De Vos, D. E., Alsters, P. L. & Feringa, B. L. Homogeneous *cis*-dihydroxylation and epoxidation of olefins with high H<sub>2</sub>O<sub>2</sub> efficiency by mixed manganese/activated carbonyl catalyst system. *Tetrahedron Lett.* **43**, 2619–2622 (2002).
40. Lee, A. W. M., Chan, W. H., Yuen, W. H., Xia, P. F. & Wong, W. Y. Ruthenium catalyzed asymmetric dihydroxylation with sultams as chiral auxiliaries. *Tetrahedron: Asymmetry* **10**,

- 1421–1424 (1999).
41. Neisius, N. M. & Plietker, B. Diastereoselective Ru-Catalyzed Cross-Metathesis-Dihydroxylation Sequence. An Efficient Approach toward Enantiomerically Enriched *syn*-Diols. *J. Org. Chem.* **73**, 3218–3227 (2008).
  42. Bhunnoo, R. A., Hu, Y., Lainé, D. I. & Brown, R. C. D. An Asymmetric Phase-Transfer Dihydroxylation Reaction. *Angew. Chem. Int. Ed.* **41**, 3479–3480 (2002).
  43. de Boer, J. W., Browne, W. R., Harutyunyan, S. R., Bini, L., Tiemersma-Wegman, T. D., Alsters, P. L., Hage, R. & Feringa, B. L. Manganese catalysed asymmetric *cis*-dihydroxylation with H<sub>2</sub>O<sub>2</sub>. *Chem. Commun.* 3747–3749 (2008).
  44. Enthaler, S., Junge, K. & Beller, M. Sustainable Metal Catalysis with Iron: From Rust to a Rising Star? *Angew. Chem. Int. Ed.* **47**, 3317–21 (2008).
  45. Chen, K. & Que Jr. *cis*-Dihydroxylation of Olefins by a Non-Heme Iron Catalyst: A Functional Model for Rieske Dioxygenases. *Angew. Chem. Int. Ed.* **38**, 2227–2229 (1999).
  46. Oldenburg, P. D., Shteinman, A. A. & Que, L. Iron-Catalyzed Olefin *cis*-Dihydroxylation Using a Bio-Inspired *N,N,O*-Ligand. *J. Am. Chem. Soc.* **127**, 15672–15673 (2005).
  47. Costas, M., Tipton, A. K., Chen, K., Jo, D.-H. & Que, L. Modeling Rieske Dioxygenases: The First Example of Iron-Catalyzed Asymmetric *cis*-Dihydroxylation of Olefins. *J. Am. Chem. Soc.* **123**, 6722–6723 (2001).
  48. Suzuki, K., Oldenburg, P. D. & Que, L. Iron-Catalyzed Asymmetric Olefin *cis*-Dihydroxylation with 97 % Enantiomeric Excess. *Angew. Chem.* **120**, 1913–1915 (2008).
  49. Wohlgemuth, R. Interfacing biocatalysis and organic synthesis. *J. Chem. Technol. Biotechnol.* **82**, 1055–1062 (2007).
  50. Hult, K. & Berglund, P. Engineered enzymes for improved organic synthesis. *Curr. Opin. Biotechnol.* **14**, 395–400 (2003).
  51. Visser, S. P. de & Kumar, D. *Iron-Containing Enzymes*. (Royal Society of Chemistry, 2011).
  52. Faber, K. *Biotransformations in Organic Chemistry*. (Springer Berlin Heidelberg, 2011).
  53. Bornscheuer, U. T. & Kazlauskas, R. J. *Hydrolases in Organic Synthesis*. (Wiley-VCH Verlag GmbH & Co. KGaA, 2005).
  54. Bellucci, G., Capitani, I., Chiappe, C. & Marioni, F. Product Enantioselectivity of the Microsomal and Cytosolic Epoxide Hydrolase catalysed Hydrolysis of *meso* Epoxides. *J. Chem. Soc., Chem. Commun.* 1170–1171 (1989).
  55. Zhao, L., Han, B., Huang, Z., Miller, M., Huang, H., Malashock, D. S., Zhu, Z., Milan, A., Robertson, D. E., Weiner, D. P. & Burk, M. J. Epoxide Hydrolase-Catalyzed Enantioselective Synthesis of Chiral 1,2-Diols via Desymmetrization of *meso*-Epoxides. *J. Am. Chem. Soc.* **126**, 11156–11157 (2004).
  56. Simon, R. C., Mutti, F. G. & Kroutil, W. Biocatalytic synthesis of enantiopure building blocks for pharmaceuticals. *Drug Discov. Today Technol.* **10**, e37–e44 (2013).
  57. Edegger, K., Stampfer, W., Seisser, B., Faber, K., Mayer, S. F., Oehrlein, R., Hafner, A. & Kroutil, W. Regio- and Stereoselective Reduction of Diketones and Oxidation of Diols by Biocatalytic Hydrogen Transfer. *Eur. J. Org. Chem.* **2006**, 1904–1909 (2006).
  58. Kulig, J., Simon, R. C., Rose, C. A., Husain, S. M., Häckh, M., Lüdeke, S., Zeitler, K., Kroutil, W., Pohl, M. & Rother, D. Stereoselective synthesis of bulky 1,2-diols with alcohol dehydrogenases. *Catal. Sci. Technol.* **2**, 1580–1589 (2012).
  59. Bortolini, O., Fantin, G., Fogagnolo, M., Giovannini, P. P., Guerrini, A. & Medici, A. An Easy Approach to the Synthesis of Optically Active *vic*-Diols: A New Single-Enzyme System. *J. Org. Chem.* **62**, 1854–1856 (1997).
  60. Lee, L. G. & Whitesides, G. M. Preparation of optically active 1,2-diols and  $\alpha$ -hydroxy ketones using glycerol dehydrogenase as catalyst. Limits to enzyme-catalyzed synthesis due to noncompetitive and mixed inhibition by product. *J. Org. Chem.* **51**, 25–36 (1986).

61. Samland, A. K. & Sprenger, G. A. Microbial aldolases as C–C bonding enzymes—unknown treasures and new developments. *Appl. Microbiol. Biotechnol.* **71**, 253–264 (2006).
62. Barry, S. M. & Challis, G. L. Mechanism and Catalytic Diversity of Rieske Non-Heme Iron-Dependent Oxygenases. *ACS Catal.* **3**, 2362–2370 (2013).
63. Ferraro, D. J., Gakhar, L. & Ramaswamy, S. Rieske business: Structure–function of Rieske non-heme oxygenases. *Biochem. Biophys. Res. Commun.* **338**, 175–190 (2005).
64. Gibson, D. T. & Parales, R. E. Aromatic hydrocarbon dioxygenases in environmental biotechnology. *Curr. Opin. Biotechnol.* **11**, 236–243 (2000).
65. Parales, R. E. & Ju, K.-S. in *Comprehensive Biotechnology* 115–134 (Elsevier, 2011).
66. Gibson, D. T., Koch, J. R. & Kallio, R. E. Oxidative Degradation of Aromatic Hydrocarbons by Microorganisms. I. Enzymic Formation of Catechol from Benzene. *Biochemistry* **7**, 2653–2662 (1968).
67. Gibson, D. T. Microbial Degradation of Aromatic Compounds. *Science* **161**, 1093–1097 (1968).
68. Fuchs, G., Boll, M. & Heider, J. Microbial degradation of aromatic compounds - from one strategy to four. *Nat. Rev. Microbiol.* **9**, 803–816 (2011).
69. Chen, J., Huang, W., Han, J. & Cao, S. The Characterization and Application of Biological Remediation Technology for Organic Contaminants. *Int. J. Environ. Res.* **5**, 515–530 (2011).
70. Nojiri, H., Habe, H. & Omori, T. Bacterial degradation of aromatic compounds via angular dioxygenation. *J. Gen. Appl. Microbiol.* **47**, 279–305 (2001).
71. Chakraborty, J., Ghosal, D., Dutta, A. & Dutta, T. K. An insight into the origin and functional evolution of bacterial aromatic ring-hydroxylating oxygenases. *J. Biomol. Struct. Dyn.* **30**, 419–436 (2012).
72. Torrelo, G., Hanefeld, U. & Hollmann, F. Biocatalysis. *Catal. Lett.* **145**, 309–345 (2015).
73. Takami, W., Yoshida, T., Nojiri, H., Yamane, H. & Omori, T. Oxidation of chlorinated olefins by *Escherichia coli* transformed with dimethyl sulfide monooxygenase genes or cumene dioxygenase genes. *J. Gen. Appl. Microbiol.* **45**, 69–75 (1999).
74. Boyd, D. R., Clarke, D., Cleij, M. C., Hamilton, J. T. G. & Sheldrake, G. N. Bacterial Biotransformation of Isoprene and Related Dienes. *Monatsh. Chem.* **131**, 673–685 (2000).
75. Boyd, D. R., Sharma, N. D., Bowers, N. I., Brannigan, I. N., Groocock, M. R., Malone, J. F., McConville, G. & Allen, C. C. R. Biocatalytic Asymmetric Dihydroxylation of Conjugated Mono- and Poly-alkenes to Yield Enantiopure Cyclic *cis*-Diols. *Adv. Synth. Catal.* **347**, 1081–1089 (2005).
76. Mason, J. R. & Cammack, R. The Electron-Transport Proteins of Hydroxylating Bacterial Dioxygenases. *Annu. Rev. Microbiol.* **46**, 277–305 (1992).
77. Costas, M., Mehn, M. P., Jensen, M. P. & Que, L. Dioxygen Activation at Mononuclear Nonheme Iron Active Sites: Enzymes, Models, and Intermediates. *Chem. Rev.* **104**, 939–986 (2004).
78. Hegg, E. L. & Que, L. The 2-His-1-carboxylate facial triad - An emerging structural motif in mononuclear non-heme iron(II) enzymes. *Eur. J. Biochem.* **250**, 625–629 (1997).
79. Koehntop, K. D., Emerson, J. P. & Que, L. The 2-His-1-carboxylate facial triad: a versatile platform for dioxygen activation by mononuclear non-heme iron(II) enzymes. *J. Biol. Inorg. Chem.* **10**, 87–93 (2005).
80. Straganz, G. D. & Nidetzky, B. Variations of the 2-His-1-carboxylate Theme in Mononuclear Non-Heme Fe<sup>II</sup> Oxygenases. *ChemBioChem* **7**, 1536–1548 (2006).
81. Shergill, J. K., Joannou, C. L., Mason, J. R. & Cammack, R. Coordination of the Rieske-Type [2Fe-2S] Cluster of the Terminal Iron-Sulfur Protein of *Pseudomonas putida* Benzene 1,2-Dioxygenase, Studied by One- and Two-Dimensional Electron Spin-Echo Envelope Modulation Spectroscopy. *Biochemistry* **34**, 16533–16542 (1995).
82. Parales, R. E. & Resnick, S. M. in *Biocatalysis in the Pharmaceutical and Biotechnology*

- Industries* 299–331 (CRC Press, 2006).
83. Sydor, P. K., Barry, S. M., Odulate, O. M., Barona-Gomez, F., Haynes, S. W., Corre, C., Song, L. & Challis, G. L. Regio- and stereodivergent antibiotic oxidative carbocyclizations catalysed by Rieske oxygenase-like enzymes. *Nat. Chem.* **3**, 388–392 (2011).
  84. Han, J., Kim, S.-Y., Jung, J., Lim, Y., Ahn, J.-H., Kim, S.-I. & Hur, H.-G. Epoxide Formation on the Aromatic B Ring of Flavanone by Biphenyl Dioxygenase of *Pseudomonas pseudoalcaligenes* KF707. *Appl. Environ. Microbiol.* **71**, 5354–5361 (2005).
  85. Batie, C. I., Ballou, D. P. & Correll, C. C. in *Chemistry and Biochemistry of Flavoenzymes Volume III* 543–556 (CRC Press, 1992).
  86. Werlen, C., Kohler, H. P. & van der Meer, J. R. The Broad Substrate Chlorobenzene Dioxygenase and *cis*-Chlorobenzene Dihydrodiol Dehydrogenase of *Pseudomonas* sp. Strain P51 Are Linked Evolutionarily to the Enzymes for Benzene and Toluene Degradation. *J. Biol. Chem.* **271**, 4009–4016 (1996).
  87. Kweon, O., Kim, S.-J., Baek, S., Chae, J.-C., Adjei, M. D., Baek, D.-H., Kim, Y.-C. & Cerniglia, C. E. A new classification system for bacterial Rieske non-heme iron aromatic ring-hydroxylating oxygenases. *BMC Biochem.* **9**, 11 (2008).
  88. Nam, J.-W., Nojiri, H., Yoshida, T., Habe, H., Yamane, H. & Omori, T. New Classification System for Oxygenase Components Involved in Ring-Hydroxylating Oxygenations. *Biosci. Biotechnol. Biochem.* **65**, 254–263 (2001).
  89. Joern, J. M., Sakamoto, T., Arisawa, A. & Arnold, F. H. A Versatile High Throughput Screen for Dioxygenase Activity Using Solid-Phase Digital Imaging. *J. Biomol. Screen.* **6**, 219–223 (2001).
  90. Wackett, L. P., Kwart, L. D. & Gibson, D. T. Benzylic Monooxygenation Catalyzed by Toluene Dioxygenase from *Pseudomonas putida*. *Biochemistry* **27**, 1360–1367 (1988).
  91. Boyd, D. R., McMordie, R. A. S., Sharma, N. D., Dalton, H., Williams, P. & Jenkins, R. O. Stereospecific Benzylic Hydroxylation of Bicyclic Alkenes by *Pseudomonas putida*: Isolation of (+)-*R*-1-Hydroxy-1,2-dihydronaphthalene, an Arene Hydrate of Naphthalene from Metabolism of 1,2-Dihydronaphthalene. *J. Chem. Soc., Chem. Commun.* 339–340 (1989).
  92. Gibson, D. T., Resnick, S. M., Lee, K., Brand, J. M., Torok, D. S., Wackett, L. P., Schocken, M. J. & Haigler, B. E. Desaturation, Dioxygenation, and Monooxygenation Reactions Catalyzed by Naphthalene Dioxygenase from *Pseudomonas* sp. Strain 9816-4. *J. Bacteriol.* **177**, 2615–2621 (1995).
  93. Resnick, S., Lee, K. & Gibson, D. Diverse reactions catalyzed by naphthalene dioxygenase from *Pseudomonas* sp strain NCIB 9816. *J. Ind. Microbiol. Biotechnol.* **17**, 438–457 (1996).
  94. Mugford, P. F., Wagner, U. G., Jiang, Y., Faber, K. & Kazlauskas, R. J. Enantiocomplementary Enzymes: Classification, Molecular Basis for Their Enantioselectivity, and Prospects for Mirror-Image Biotransformations. *Angew. Chem. Int. Ed.* **47**, 8782–8793 (2008).
  95. Boyd, D. R., Sharma, N. D., Boyle, R., McMurray, B. T., Evans, T. A., Malone, J. F., Dalton, H., Chima, J. & Sheldrake, G. N. Biotransformation of Unsaturated Heterocyclic Rings by *Pseudomonas putida* to Yield *cis*-Diols. *J. Chem. Soc., Chem. Commun.* 49–51 (1993).
  96. Modyanova, L. & Azerad, R. Dioxygenase-catalysed formation of dihydrodiol metabolites of *N*-methyl-2-pyridone. *Tetrahedron Lett.* **41**, 3865–3869 (2000).
  97. Bowers, N. I., Boyd, D. R., Sharma, N. D., Kennedy, M. A., Sheldrake, G. N. & Dalton, H. Stereoselective *cis*-dihydroxylation of azulene and related non-aromatic polyenes. *Tetrahedron: Asymmetry* **9**, 1831–1834 (1998).
  98. Lange, C. C. & Wackett, L. P. Oxidation of Aliphatic Olefins by Toluene Dioxygenase: Enzyme Rates and Product Identification. *J. Bacteriol.* **179**, 3858–3865 (1997).
  99. Hurtubise, Y., Barriault, D. & Sylvestre, M. Involvement of the Terminal Oxygenase  $\beta$  Subunit in the Biphenyl Dioxygenase Reactivity Pattern toward Chlorobiphenyls. *J. Bacteriol.* **180**, 5828–5835 (1998).
  100. Bassan, A., Blomberg, M. R. A., Borowski, T. & Siegbahn, P. E. M. Oxygen Activation by Rieske Non-Heme Iron Oxygenases, a Theoretical Insight. *J. Phys. Chem. B* **108**, 13031–

- 13041 (2004).
101. Karlsson, A., Parales, J. V., Parales, R. E., Gibson, D. T., Eklund, H. & Ramaswamy, S. Crystal Structure of Naphthalene Dioxygenase: Side-on Binding of Dioxygen to Iron. *Science* **299**, 1039–1042 (2003).
  102. Kauppi, B., Lee, K., Carredano, E., Parales, R. E., Gibson, D. T., Eklund, H. & Ramaswamy, S. Structure of an aromatic-ring-hydroxylating dioxygenase-naphthalene 1,2-dioxygenase. *Structure* **6**, 571–586 (1998).
  103. Ohta, T., Chakrabarty, S., Lipscomb, J. D. & Solomon, E. I. Near-IR MCD of the Nonheme Ferrous Active Site in Naphthalene 1,2-Dioxygenase: Correlation to Crystallography and Structural Insight into the Mechanism of Rieske Dioxygenases. *J. Am. Chem. Soc.* **130**, 1601–1610 (2008).
  104. Parales, R. E., Parales, J. V. & Gibson, D. T. Aspartate 205 in the Catalytic Domain of Naphthalene Dioxygenase Is Essential for Activity. *J. Bacteriol.* **181**, 1831–1837 (1999).
  105. Wolfe, M. D., Parales, J. V., Gibson, D. T. & Lipscomb, J. D. Single Turnover Chemistry and Regulation of O<sub>2</sub> Activation by the Oxygenase Component of Naphthalene 1,2-Dioxygenase. *J. Biol. Chem.* **276**, 1945–1953 (2001).
  106. Chakrabarty, S., Austin, R. N., Deng, D., Groves, J. T. & Lipscomb, J. D. Radical Intermediates in Monooxygenase Reactions of Rieske Dioxygenases. *J. Am. Chem. Soc.* **129**, 3514–3515 (2007).
  107. Schrittwieser, J. H. & Resch, V. The role of biocatalysis in the asymmetric synthesis of alkaloids. *RSC Adv.* **3**, 17602–17632 (2013).
  108. Duchek, J., Adams, D. R. & Hudlicky, T. Chemoenzymatic Synthesis of Inositols, Conduritols, and Cycylitol Analogues. *Chem. Rev.* **111**, 4223–4258 (2011).
  109. Newman, L. M., Garcia, H., Hudlicky, T. & Selifonov, S. A. Directed evolution of the dioxygenase complex for the synthesis of furanone flavor compounds. *Tetrahedron* **60**, 729–734 (2004).
  110. Hudlicky, T., Luna, H., Barbiera, G. & Kwart, L. D. Enantioselective Synthesis through Microbial Oxidation of Arenes. 1. Efficient Preparation of Terpene and Prostanoid Synthons. *J. Am. Chem. Soc.* **110**, 4735–4741 (1988).
  111. Hudlicky, T. Chemoenzymatic synthesis of complex natural and unnatural products: morphine, pancratistatin, and their analogs. *Arkivoc* **2006**, 276–291 (2006).
  112. Akgün, H. & Hudlicky, T. Total Syntheses of *ent*-Conduramine A and *ent*-7-Deoxypancratistatin. *Tetrahedron Lett.* **40**, 3081–3084 (1999).
  113. Ballard, D. G. H., Courtis, A., Shirley, I. M. & Taylor, S. C. A biotech route to polyphenylene. *J. Chem. Soc., Chem. Commun.* 954–955 (1983).
  114. Ensley, B., Ratzkin, B., Osslund, T., Simon, M., Wackett, L. & Gibson, D. Expression of Naphthalene Oxidation Genes in *Escherichia coli* Results in the Biosynthesis of Indigo. *Science* **222**, 167–169 (1983).
  115. Furukawa, K. Engineering dioxygenases for efficient degradation of environmental pollutants. *Curr. Opin. Biotechnol.* **11**, 244–249 (2000).
  116. Behrens, G. A., Hummel, A., Padhi, S. K., Schätzle, S. & Bornscheuer, U. T. Discovery and Protein Engineering of Biocatalysts for Organic Synthesis. *Adv. Synth. Catal.* **353**, 2191–2215 (2011).
  117. Hibbert, E. G., Baganz, F., Hailes, H. C., Ward, J. M., Lye, G. J., Woodley, J. M. & Dalby, P. A. Directed evolution of biocatalytic processes. *Biomol. Eng.* **22**, 11–19 (2005).
  118. Davids, T., Schmidt, M., Böttcher, D. & Bornscheuer, U. T. Strategies for the discovery and engineering of enzymes for biocatalysis. *Curr. Opin. Chem. Biol.* **17**, 215–220 (2013).
  119. Dror, A. & Fishman, A. Engineering Non-Heme Mono- and Dioxygenases for Biocatalysis. *Comput. Struct. Biotechnol. J.* **2**, e201209011 (2012).

120. Steiner, K. & Schwab, H. Recent Advances in Rational Approaches for Enzyme Engineering. *Comput. Struct. Biotechnol. J.* **2**, e201209010 (2012).
121. Reetz, M. T. The Importance of Additive and Non-Additive Mutational Effects in Protein Engineering. *Angew. Chem. Int. Ed.* **52**, 2658–2666 (2013).
122. Chica, R. A., Doucet, N. & Pelletier, J. N. Semi-rational approaches to engineering enzyme activity: combining the benefits of directed evolution and rational design. *Curr. Opin. Biotechnol.* **16**, 378–384 (2005).
123. Morley, K. L. & Kazlauskas, R. J. Improving enzyme properties: when are closer mutations better? *Trends Biotechnol.* **23**, 231–237 (2005).
124. Parales, R. E., Resnick, S. M., Yu, C.-L., Boyd, D. R., Sharma, N. D. & Gibson, D. T. Regioselectivity and Enantioselectivity of Naphthalene Dioxygenase during Arene *cis*-Dihydroxylation: Control by Phenylalanine 352 in the  $\alpha$  Subunit. *J. Bacteriol.* **182**, 5495–5504 (2000).
125. Yu, C. L., Parales, R. E. & Gibson, D. T. Multiple mutations at the active site of naphthalene dioxygenase affect regioselectivity and enantioselectivity. *J. Ind. Microbiol. Biotechnol.* **27**, 94–103 (2001).
126. Parales, R. E. The role of active-site residues in naphthalene dioxygenase. *J. Ind. Microbiol. Biotechnol.* **30**, 271–278 (2003).
127. Pollmann, K. Rational engineering of the regioselectivity of TecA tetrachlorobenzene dioxygenase for the transformation of chlorinated toluenes. *Microbiology* **149**, 903–913 (2003).
128. Suenaga, H., Watanabe, T., Sato, M., Ngadiman & Furukawa, K. Alteration of Regiospecificity in Biphenyl Dioxygenase by Active-Site Engineering. *J. Bacteriol.* **184**, 3682–3688 (2002).
129. Erickson, B. D. & Mondello, F. J. Enhanced Biodegradation of Polychlorinated Biphenyls after Site-Directed Mutagenesis of a Biphenyl Dioxygenase Gene. *Appl. Environ. Microbiol.* **59**, 3858–3862 (1993).
130. Kaur, J. & Sharma, R. Directed Evolution: An Approach to Engineer Enzymes. *Crit. Rev. Biotechnol.* **26**, 165–199 (2006).
131. Zhao, H., Chockalingam, K. & Chen, Z. Directed evolution of enzymes and pathways for industrial biocatalysis. *Curr. Opin. Biotechnol.* **13**, 104–110 (2002).
132. Stebel, S. C., Gaida, A., Arndt, K. M. & Müller, K. M. in *Molecular Biomechanics Handbook* 631–656 (Humana Press, 2008).
133. Brühlmann, F. & Chen, W. Tuning Biphenyl Dioxygenase for Extended Substrate Specificity. *Biotechnol. Bioeng.* **63**, 544–551 (1999).
134. Suenaga, H., Goto, M. & Furukawa, K. Emergence of Multifunctional Oxygenase Activities by Random Priming Recombination. *J. Biol. Chem.* **276**, 22500–22506 (2001).
135. Kumamaru, T., Suenaga, H., Mitsuoka, M., Watanabe, T. & Furukawa, K. Enhanced degradation of polychlorinated biphenyls by directed evolution of biphenyl dioxygenase. *Nat. Biotechnol.* **16**, 663–666 (1998).
136. Kimura, N., Nishi, A., Goto, M. & Furukawa, K. Functional Analyses of a Variety of Chimeric Dioxygenases Constructed from Two Biphenyl Dioxygenases That Are Similar Structurally but Different Functionally. *J. Bacteriol.* **179**, 3936–3943 (1997).
137. Ang, E. L., Obbard, J. P. & Zhao, H. Directed evolution of aniline dioxygenase for enhanced bioremediation of aromatic amines. *Appl. Microbiol. Biotechnol.* **81**, 1063–1070 (2009).
138. Sakamoto, T., Joern, J. M., Arisawa, A. & Arnold, F. H. Laboratory Evolution of Toluene Dioxygenase To Accept 4-Picoline as a Substrate. *Appl. Environ. Microbiol.* **67**, 3882–3887 (2001).
139. Johannes, T. W., Woodyer, R. D. & Zhao, H. in *Enzyme Assays* 77–93 (Wiley-VCH Verlag GmbH & Co. KGaA, 2006).
140. Reetz, M. T., Kahakeaw, D. & Lohmer, R. Addressing the Numbers Problem in Directed

- Evolution. *ChemBioChem* **9**, 1797–1804 (2008).
141. Lutz, S. Beyond directed evolution—semi-rational protein engineering and design. *Curr. Opin. Biotechnol.* **21**, 734–743 (2010).
  142. Shainsky, J., Bernath-Levin, K., Isaschar-Ovdat, S., Glaser, F. & Fishman, A. Protein engineering of nirobenzene dioxygenase for enantioselective synthesis of chiral sulfoxides. *Protein Eng. Des. Sel.* **26**, 335–345 (2013).
  143. Kagami, O., Shindo, K., Kyojima, A., Takeda, K., Ikenaga, H., Furukawa, K. & Misawa, N. Protein Engineering on Biphenyl Dioxygenase for Conferring Activity to Convert 7-Hydroxyflavone and 5,7-Dihydroxyflavone (Chrysin). *J. Biosci. Bioeng.* **106**, 121–127 (2008).
  144. Koeller, K. M. & Wong, C.-H. Enzymes for chemical synthesis. *Nature* **409**, 232–240 (2001).
  145. Parales, R. E., Lee, K., Resnick, S. M., Jiang, H., Lessner, D. J. & Gibson, D. T. Substrate Specificity of Naphthalene Dioxygenase: Effect of Specific Amino Acids at the Active Site of the Enzyme. *J. Bacteriol.* **182**, 1641–1649 (2000).
  146. Swift, R. J., Carter, S. F., Widdowson, D. A., Mason, J. R. & Leak, D. J. Expression of benzene dioxygenase from *Pseudomonas putida* ML2 in *cis*-1,2-cyclohexanediol-degrading pseudomonads. *Appl. Microbiol. Biotechnol.* **55**, 721–726 (2001).
  147. Aoki, H., Kimura, T., Habe, H., Yamane, H., Kodama, T. & Omori, T. Cloning, Nucleotide Sequence, and Characterization of the Genes Encoding Enzymes Involved in the Degradation of Cumene to 2-Hydroxy-6-Oxo-7-Methylocta-2,4-Dienoic Acid in *Pseudomonas fluorescens* IP01. *J. Ferment. Bioeng.* **81**, 187–196 (1996).
  148. Mason, J. R., Butler, C. S., Cammack, R. & Shergill, J. K. Structural studies on the catalytic component of benzene dioxygenase from *Pseudomonas putida*. *Biochem. Soc. Trans.* **25**, 90–95 (1997).
  149. Korbie, D. J. & Mattick, J. S. Touchdown PCR for increased specificity and sensitivity in PCR amplification. *Nat. Protoc.* **3**, 1452–1456 (2008).
  150. Papworth, C., Greener, A. & Braman, J. Highly Efficient Double-Stranded, Site-Directed Mutagenesis with the Chameleon Kit. *Strateg. Mol. Biol.* **7**, 38–40 (1996).
  151. Mitchell, L. A., Cai, Y., Taylor, M., Noronha, A. M., Chuang, J., Dai, L. & Boeke, J. D. Multichange Isothermal Mutagenesis: A New Strategy for Multiple Site-Directed Mutations in Plasmid DNA. *ACS Synth. Biol.* **2**, 473–477 (2013).
  152. Hanahan, D. Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* **166**, 557–580 (1983).
  153. Bagneris, C., Cammack, R. & Mason, J. R. Subtle Difference between Benzene and Toluene Dioxygenases of *Pseudomonas putida*. *Appl. Environ. Microbiol.* **71**, 1570–1580 (2005).
  154. Yildirim, S., Zezula, J., Hudlicky, T., Witholt, B. & Schmid, A. Asymmetric Dihydroxylation of Cinnamionitrile to *trans*-3-[(5*S*,6*R*)-5,6-Dihydroxycyclohexa-1,3-dienyl]-acrylonitrile using Chlorobenzene Dioxygenase in *Escherichia coli* (pTEZ30). *Adv. Synth. Catal.* **346**, 933–942 (2004).
  155. Buhler, B., Witholt, B., Hauer, B. & Schmid, A. Characterization and Application of Xylene Monooxygenase for Multistep Biocatalysis. *Appl. Environ. Microbiol.* **68**, 560–568 (2002).
  156. Espina, L., Gelaw, T. K., de Lamo-Castellví, S., Pagán, R. & García-Gonzalo, D. Mechanism of Bacterial Inactivation by (+)-Limonene and Its Potential Use in Food Preservation Combined Processes. *PLoS One* **8**, e56769 (2013).
  157. Onaran, M. B. & Seto, C. T. Using a Lipase as a High-Throughput Screening Method for Measuring the Enantiomeric Excess of Allylic Acetates. *J. Org. Chem.* **68**, 8136–8141 (2003).
  158. Scanlon, J. T. & Willis, D. E. Calculation of Flame Ionization Detector Relative Response Factors Using the Effective Carbon Number Concept. *J. Chromatogr. Sci.* **23**, 333–340 (1985).
  159. Robert L. Grob, E. F. B. *Modern Practice of Gas Chromatography, Fourth Edition.* (Wiley-VCH Verlag GmbH & Co. KGaA, 2004).

160. Wynberg, H. & Marsman, B. Synthesis of Optically Active 2,3-Epoxycyclohexanone and the Determination of Its Absolute Configuration. *J. Org. Chem.* **45**, 158–161 (1980).
161. Granger, R., Passet, J. & Girard, J. Methyl-2 methylene-6 octadiene-2,7 ol isole de *Thymus vulgaris*. *Phytochemistry* **11**, 2301–2305 (1972).
162. Fauchet, V., Miguel, B., Taran, M. & Delmond, B. Epoxides from Myrcene: Selective Obtention. *Synth. Commun.* **23**, 2503–2510 (1993).
163. Erdik, E., Kähya, D. & Daşkapan, T. Accelerating Ligands for Osmium Tetraoxide Catalyzed Racemic Dihydroxylation of  $\alpha$ -Pinene. *Synth. Commun.* **28**, 1–7 (1998).
164. Thomas, A. F. & Bucher, W. Menthatrienes and the Oxidation of Limonene. *Helv. Chim. Acta* **53**, 770–775 (1970).
165. Dhulut, S., Bourin, A., Lannou, M.-I., Fleury, E., Lensen, N., Chelain, E., Pancrazi, A., Ardisson, J. & Fahy, J. Cyclic Allyl Carbamates in Stereoselective  $syn\ Se'$  Processes: Synthetic Approach to Sarcodictyins and Eleutherobin. *Eur. J. Org. Chem.* **2007**, 5235–5243 (2007).
166. Rico, R. & Bermejo, F. Total Synthesis of (–)-Ampullicin and (+)-Isoampullicin Two Growth Regulators from *Ampulliferina* Sp. No. 27. *Tetrahedron Lett.* **36**, 7889–7892 (1995).
167. Duan, Y., Wu, C., Chowdhury, S., Lee, M. C., Xiong, G., Zhang, W., Yang, R., Cieplak, P., Luo, R., Lee, T., Caldwell, J., Wang, J. & Kollman, P. A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations. *J. Comput. Chem.* **24**, 1999–2012 (2003).
168. Trott, O. & Olson, A. J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **31**, 455–461 (2009).
169. Lee, K. & Gibson, D. T. Stereospecific Dihydroxylation of the Styrene Vinyl Group by Purified Naphthalene Dioxygenase from *Pseudomonas* sp. Strain NCIB 9816-4. *J. Bacteriol.* **178**, 3353–3356 (1996).
170. Dong, X., Fushinobu, S., Fukuda, E., Terada, T., Nakamura, S., Shimizu, K., Nojiri, H., Omori, T., Shoun, H. & Wakagi, T. Crystal Structure of the Terminal Oxygenase Component of Cumene Dioxygenase from *Pseudomonas fluorescens* IP01. *J. Bacteriol.* **187**, 2483–2490 (2005).
171. Friemann, R., Lee, K., Brown, E. N., Gibson, D. T., Eklund, H. & Ramaswamy, S. Structures of the multicomponent Rieske non-heme iron toluene 2,3-dioxygenase enzyme system. *Acta Crystallogr. D Biol. Crystallogr.* **65**, 24–33 (2009).
172. Sauber, K., Fröhner, C., Rosenberg, G., Eberspächer, J. & Lingens, F. Purification and Properties of Pyrazon Dioxygenase from Pyrazon-Degrading Bacteria. *Eur. J. Biochem.* **74**, 89–97 (1977).
173. Catterall, G. F. & Williams, P. A. Some Properties of the Naphthalene Oxygenase from *Pseudomonas* sp. NCIB 9816. *J. Gen. Microbiol.* **67**, 117–124 (1971).
174. Seo, J., Kang, S.-I., Ryu, J.-Y., Lee, Y.-J., Park, K. D., Kim, M., Won, D., Park, H.-Y., Ahn, J.-H., Chong, Y., Kanaly, R. A., Han, J. & Hur, H.-G. Location of flavone B-ring controls regioselectivity and stereoselectivity of naphthalene dioxygenase from *Pseudomonas* sp. strain NCIB 9816-4. *Appl. Microbiol. Biotechnol.* **86**, 1451–1462 (2010).
175. Torok, D. S., Resnick, S. M., Brand, J. M., Cruden, D. L. & Gibson, D. T. Desaturation and Oxygenation of 1,2-Dihydronaphthalene by Toluene and Naphthalene Dioxygenase. *J. Bacteriol.* **177**, 5799–5805 (1995).
176. Becker, H., King, S. B., Taniguchi, M., Vanhessche, K. P. M. & Sharpless, K. B. New Ligands and Improved Enantioselectivities for the Asymmetric Dihydroxylation of Olefins. *J. Org. Chem.* **60**, 3940–3941 (1995).
177. Becker, H. & Sharpless, K. B. Eine neue Ligandenklasse für die asymmetrische Dihydroxylierung von Olefinen. *Angew. Chem.* **108**, 447–449 (1996).
178. de Carvalho, C. C. C. R. & da Fonseca, M. M. R. Biotransformation of terpenes. *Biotechnol. Adv.* **24**, 134–142 (2006).

179. Pal, D. & Chakrabarti, P. Non-hydrogen Bond Interactions Involving the Methionine Sulfur Atom. *J. Biomol. Struct. Dyn.* **19**, 115–128 (2001).
180. Zauhar, R. J., Colbert, C. L., Morgan, R. S. & Welsh, W. J. Evidence for a Strong Sulfur-Aromatic Interaction Derived from Crystallographic Data. *Biopolymers* **53**, 233–248 (2000).
181. Hudlicky, T., Boros, E. E. & Boros, C. H. New Diol Metabolites Derived by Biooxidation of Chlorostyrenes with *Pseudomonas putida*: Determination of Absolute Stereochemistry and Enantiomeric Excess by Convergent Syntheses. *Tetrahedron: Asymmetry* **4**, 1365–1386 (1993).
182. Beil, S., Mason, J. R., Timmis, K. N. & Pieper, D. H. Identification of Chlorobenzene Dioxygenase Sequence Elements Involved in Dechlorination of 1,2,4,5-Tetrachlorobenzene. *J. Bacteriol.* **180**, 5520–5528 (1998).
183. Bordoli, L., Kiefer, F., Arnold, K., Benkert, P., Battey, J. & Schwede, T. Protein structure homology modeling using SWISS-MODEL workspace. *Nat. Protoc.* **4**, 1–13 (2008).
184. Pavelka, A., Chovancova, E. & Damborsky, J. HotSpot Wizard: a web server for identification of hot spots in protein engineering. *Nucleic Acids Res.* **37**, W376–W383 (2009).
185. Parales, R. E., Emig, M. D., Lynch, N. A. & Gibson, D. T. Substrate Specificities of Hybrid Naphthalene and 2,4-Dinitrotoluene Dioxygenase Enzyme Systems. *J. Bacteriol.* **180**, 2337–2344 (1998).
186. Bühler, B., Bollhalder, I., Hauer, B., Witholt, B. & Schmid, A. Chemical Biotechnology for the Specific Oxyfunctionalization of Hydrocarbons on a Technical Scale. *Biotechnol. Bioeng.* **82**, 833–842 (2003).
187. Lopez-Romero, J. C., González-Ríos, H., Borges, A. & Simões, M. Antibacterial Effects and Mode of Action of Selected Essential Oils Components against *Escherichia coli* and *Staphylococcus aureus*. *Evid. Based Complement. Alternat. Med.* **2015**, 1–9 (2015).
188. Eiteman, M. A. & Altman, E. Overcoming acetate in *Escherichia coli* recombinant protein fermentations. *Trends Biotechnol.* **24**, 530–536 (2006).
189. Tan, H. M., Tang, H. Y., Joannou, C. L., Abdel-Wahab, N. H. & Mason, J. R. The *Pseudomonas putida* ML2 plasmid-encoded genes for benzene dioxygenase are unusual in codon usage and low in G+C content. *Gene* **130**, 33–39 (1993).
190. Habe, H., Kasuga, K., Nojiri, H., Yamane, H. & Omori, T. Analysis of Cumene (Isopropylbenzene) Degradation Genes from *Pseudomonas fluorescens* IP01. *Appl. Environ. Microbiol.* **62**, 4471–4477 (1996).
191. Jeffrey, A. M., Yeh, H. J. C., Jerina, D. M., Patel, T. R., Davey, J. F. & Gibson, D. T. Initial Reactions in the Oxidation of Naphthalene by *Pseudomonas putida*. *Biochemistry* **14**, 575–584 (1975).
192. Munro, A. W., Girvan, H. M. & McLean, K. J. Cytochrome P450–redox partner fusion enzymes. *Biochim. Biophys. Acta* **1770**, 345–359 (2007).
193. Boyd, D. R., Sharma, N. D., Haughey, S. A., Kennedy, M. A., McMurray, B. T., Sheldrake, G. N., Allen, C. C. R., Dalton, H. & Sproule, K. Toluene and naphthalene dioxygenase-catalysed sulfoxidation of alkyl aryl sulfides. *J. Chem. Soc. Perkin Trans. 1* 1929–1934 (1998).
194. Bowers, N. I., Boyd, D. R., Sharma, N. D., Goodrich, P. A., Groocock, M. R., Blacker, A. J., Goode, P. & Dalton, H. Stereoselective benzylic hydroxylation of 2-substituted indanes using toluene dioxygenase as biocatalyst. *J. Chem. Soc. Perkin Trans. 1* 1453–1462 (1999).
195. Boyd, D. R., Sharma, N. D., Bowers, N. I., Boyle, R., Harrison, J. S., Lee, K., Bugg, T. D. H. & Gibson, D. T. Stereochemical and mechanistic aspects of dioxygenase-catalysed benzylic hydroxylation of indene and chromane substrates. *Org. Biomol. Chem.* **1**, 1298–1307 (2003).
196. Paramesvaran, J., Hibbert, E. G., Russell, A. J. & Dalby, P. A. Distributions of enzyme residues yielding mutants with improved substrate specificities from two different directed evolution strategies. *Protein Eng. Des. Sel.* **22**, 401–411 (2009).
197. Witzig, R., Junca, H., Hecht, H.-J. & Pieper, D. H. Assessment of Toluene/Biphenyl Dioxygenase Gene Diversity in Benzene-Polluted Soils: Links between Benzene

- Biodegradation and Genes Similar to Those Encoding Isopropylbenzene Dioxygenases. *Appl. Environ. Microbiol.* **72**, 3504–3514 (2006).
198. Schalk, M. & Croteau, R. A single amino acid substitution (F363I) converts the regiochemistry of the spearmint (-)-limonene hydroxylase from a C6- to a C3-hydroxylase. *Proc. Natl. Acad. Sci.* **97**, 11948–11953 (2000).
199. Seo, J., Kang, S.-I., Kim, M., Han, J. & Hur, H.-G. Flavonoids biotransformation by bacterial non-heme dioxygenases, biphenyl and naphthalene dioxygenase. *Appl. Microbiol. Biotechnol.* **91**, 219–228 (2011).
200. Hobuß, D., Thöne, C., Laschat, S. & Baro, A. Synthesis of Novel Chiral Bisphosphinites from  $\alpha$ -Pinene. *Synthesis* **4**, 2053–2056 (2003).
201. Lee, K. Benzene-Induced Uncoupling of Naphthalene Dioxygenase Activity and Enzyme Inactivation by Production of Hydrogen Peroxide. *J. Bacteriol.* **181**, 2719–2725 (1999).
202. Fluxá, V. S., Wahler, D. & Reymond, J.-L. Enzyme assay and activity fingerprinting of hydrolases with the red-chromogenic adrenaline test. *Nat. Protoc.* **3**, 1270–1277 (2008).
203. Collins, A. M., Woodley, J. M. & Liddell, J. M. Determination of reactor operation for the microbial hydroxylation of toluene in a two-liquid phase process. *J. Ind. Microbiol.* **14**, 382–388 (1995).
204. Malca, S. H. Substrate characterization and protein engineering of bacterial cytochrome P450 monooxygenases for the bio-based synthesis of omega-hydroxy aliphatic compounds. (Universität Stuttgart, 2013).
205. van Beilen, J. B., Duetz, W. A., Schmid, A. & Witholt, B. Practical issues in the application of oxygenases. *Trends Biotechnol.* **21**, 170–177 (2003).
206. Wriessnegger, T., Augustin, P., Engleder, M., Leitner, E., Müller, M., Kaluzna, I., Schürmann, M., Mink, D., Zellnig, G., Schwab, H. & Pichler, H. Production of the sesquiterpenoid (+)-nootkatone by metabolic engineering of *Pichia pastoris*. *Metab. Eng.* **24**, 18–29 (2014).
207. Lynch, R. M., Woodley, J. M. & Lilly, M. D. Process design for the oxidation of fluorobenzene to fluorocatechol by *Pseudomonas putida*. *J. Biotechnol.* **58**, 167–175 (1997).
208. Reddy, J., Lee, C., Neeper, M., Greasham, R. & Zhang, J. Development of a bioconversion process for production of *cis*-1S,2*R*-indandiol from indene by recombinant *Escherichia coli* constructs. *Appl. Microbiol. Biotechnol.* **51**, 614–620 (1999).
209. Becker, H., Soler, M. A. & Barry Sharpless, K. Selective Asymmetric Dihydroxylation of Polyenes. *Tetrahedron* **51**, 1345–1376 (1995).
210. Sundermeier, U., Döbler, C. & Beller, M. in *Modern Oxidation Methods* **1**, 1–20 (Wiley-VCH Verlag GmbH & Co. KGaA, 2005).
211. Constable, D. J. C., Dunn, P. J., Hayler, J. D., Humphrey, G. R., Leazer, Jr., J. L., Linderman, R. J., Lorenz, K., Manley, J., Pearlman, B. A., Wells, A., Zaks, A. & Zhang, T. Y. Key green chemistry research areas - a perspective from pharmaceutical manufacturers. *Green Chem.* **9**, 411–420 (2007).
212. Duetz, W. A., Bouwmeester, H., Beilen, J. B. & Witholt, B. Biotransformation of limonene by bacteria, fungi, yeasts, and plants. *Appl. Microbiol. Biotechnol.* **61**, 269–277 (2003).
213. Młochowski, J. & Wójtowicz-Młochowska, H. Developments in Synthetic Application of Selenium(IV) Oxide and Organoselenium Compounds as Oxygen Donors and Oxygen-Transfer Agents. *Molecules* **20**, 10205–10243 (2015).
214. Reetz, M. T., Bocola, M., Carballeira, J. D., Zha, D. & Vogel, A. Expanding the Range of Substrate Acceptance of Enzymes: Combinatorial Active-Site Saturation Test. *Angew. Chem. Int. Ed.* **44**, 4192–4196 (2005).
215. Que, L. & Ho, R. Y. N. Dioxygen Activation by Enzymes with Mononuclear Non-Heme Iron Active Sites. *Chem. Rev.* **96**, 2607–2624 (1996).
216. Zhang, W., Liu, Y., Yan, J., Cao, S., Bai, F., Yang, Y., Huang, S., Yao, L., Anzai, Y., Kato, F., Podust, L. M., Sherman, D. H. & Li, S. New Reactions and Products Resulting from Alternative

- Interactions between the P450 Enzyme and Redox Partners. *J. Am. Chem. Soc.* **136**, 3640–3646 (2014).
217. Li, C., Zhang, L., Zhang, C., Hirao, H., Wu, W. & Shaik, S. Which Oxidant Is Really Responsible for Sulfur Oxidation by Cytochrome P450? *Angew. Chem. Int. Ed.* **46**, 8168–8170 (2007).
218. Scheps, D., Honda Malca, S., Richter, S. M., Marisch, K., Nestl, B. M. & Hauer, B. Synthesis of  $\omega$ -hydroxy dodecanoic acid based on an engineered CYP153A fusion construct. *Microb. Biotechnol.* **6**, 694–707 (2013).
219. Tauber, K., Fuchs, M., Sattler, J. H., Pitzer, J., Pressnitz, D., Koszelewski, D., Faber, K., Pfeffer, J., Haas, T. & Kroutil, W. Artificial Multi-Enzyme Networks for the Asymmetric Amination of *sec*-Alcohols. *Chem. Eur. J.* **19**, 4030–4035 (2013).
220. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J. & Higgins, D. G. Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).
221. Rice, P., Longden, I. & Bleasby, A. EMBOSS: The European Molecular Biology Open Software Suite. *Trends Genet.* **16**, 276–277 (2000).
222. Allen, C. C. R. in *Science of Synthesis: Biocatalysis in Organic Synthesis Vol. 3* 1–20 (Georg Thieme Verlag, 2015).
223. Reetz, M. T., Carballeira, J. D., Peyralans, J., Höbenreich, H., Maichele, A. & Vogel, A. Expanding the Substrate Scope of Enzymes: Combining Mutations Obtained by CASTing. *Chem. Eur. J.* **12**, 6031–6038 (2006).

## 7. SUPPLEMENTARY

### 7.1. Plasmids

**Table S23:** Plasmid constructs encoding for different ROs and variants thereof used in this work.

Name	Gene insert	ITB No.
pDTG141	NDO operon: oxygenase, ferredoxin and reductase	pITB1009
pDTG141_A206M	NDO operon with oxygenase variant A206M	-
pDTG141_A206I	NDO operon with oxygenase variant A206I	pITB1010
pJRM501	BDO operon: oxygenase, ferredoxin and reductase	pITB1014
pJRM501_M220A	BDO operon with oxygenase variant M220A	pITB1015
pIP107D	CDO operon: oxygenase, ferredoxin and reductase	pITB1011
pIP107D_M232A	CDO operon with oxygenase variant M232A	pITB1012
pIP107D_M232G	CDO operon with oxygenase variant M232G	pITB1253
pIP107D_M232F	CDO operon with oxygenase variant M232F	pITB1254
pIP107D_M232L	CDO operon with oxygenase variant M232L	pITB1255
pIP107D_M232I	CDO operon with oxygenase variant M232I	pITB1256
pIP107D_M232V	CDO operon with oxygenase variant M232V	pITB1257
pIP107D_M232Y	CDO operon with oxygenase variant M232Y	pITB1258
pIP107D_M232H	CDO operon with oxygenase variant M232H	pITB1259
pIP107D_M232N	CDO operon with oxygenase variant M232N	pITB1260
pIP107D_M232D	CDO operon with oxygenase variant M232D	pITB1261
pIP107D_M232C	CDO operon with oxygenase variant M232C	pITB1262
pIP107D_M232S	CDO operon with oxygenase variant M232S	pITB1263
pIP107D_F278A	CDO operon with oxygenase variant F278A	-
pIP107D_F278C	CDO operon with oxygenase variant F278C	-
pIP107D_I288V	CDO operon with oxygenase variant F288V	-
pIP107D_I288P	CDO operon with oxygenase variant I288P	-
pIP107D_I336V	CDO operon with oxygenase variant I336V	-
pIP107D_F378S	CDO operon with oxygenase variant F378S	-

## 7.2. Sequencing primers

**Table S24:** Primers used for sequencing of RO genes in the plasmid constructs. As RO operons were large in size, several primers had to be used for complete sequencing of the gene insert. fwd: forward primer, rev: reverse primer.

Primer name	Sequence
pIP107D <i>lacP</i> _fwd	5'-GGCTTTACACTTTATGCTTCCG-3'
pIP107D 1_fwd	5'-GTGACAAAAAGAGGGTGACTG-3'
pIP107D 2_fwd	5'-GGTCTTTCATCGTAGTGGATGC-3'
pIP107D 3_fwd	5'-CAGAAATGCCCGTCAGTCTTG-3'
pIP107D 4_fwd	5'-GCCAATCTCCTCGGGACTTTGC-3'
pIP107D 5_fwd	5'-GCATTTTCCGAGTGCGTAC-3'
pIP107D 6_fwd	5'-GCCTACGACCGACTGCTATTAG-3'
pIP107D 7_fwd	5'-CCATGTTAGGCAAGTCTATCCC-3'
pIP107D 8_rev	5'-GCTTTGACGTATGCGGTGTG-3'
pDTG141 T7P_fwd	5'-TAATACGACTCACTATAGGG-3'
pDTG141 1_fwd	5'-GCCGACGAAATTGCACTCAC-3'
pDTG141 2_fwd	5'-GCGATGGTTGAAGCGTTG-3'
pDTG141 3_fwd	5'-GGTAAGTGAATCTGGTCTGAGC-3'
pDTG141 4_fwd	5'-GCAAATGACCTCCAAATACGG-3'
pDTG141 5_fwd	5'-CTC CAA CTGGGCTGAGTTC-3'
pDTG141 6_fwd	5'-AAGCATTGGTAACTGTCAGACC-3'
pJRM501 pKK223-3_fwd	5'-CGACATCATAAC GGTTCTGGC-3'
pJRM501 1_fwd	5'-GAAGACCAGTGAGATAGAAACG-3'
pJRM501 2_fwd	5'-GGAAATTCGCAGCAGAGCAG-3'
pJRM501 3_fwd	5'-GCTGAAAATGATGACCTCACC-3'
pJRM501 4_fwd	5'-CGCAATCGCTTGGAAAGACAAC-3'
pJRM501 5_fwd	5'-CTTCCCTATGACCGACCATCC-3'
pJRM501 6_fwd	5'-GTTTGGAATGTGATCGTGGTG-3'
pJRM501 7_fwd	5'-CAGAAAAGCTGGCAGATGTGC-3'

7.3. Genes

**NDO gene cluster** (*nahAa* → *nahAb* → *nahAc* → *nahAd*)

NCBI reference sequence:NC\_004999.1

ATGGAACCTCCTATACAACCGAACAAATCGCATAATTCCTTCAGTGCCGGTGCCAACCTTCTGGAAGTGCTTC  
 GCGAGAACGGTGTAGCTATTTCTACAGTTGCTGTCTGGCGTTCGCGAACCTGTGCTGCCGGGTTATAG  
 ATGGCAGTGTCAATTGATTCTGGGGCGGAAAATGGGCAATCAAACCTACCAGCAACAGCATGTGCTGCCT  
 GTCAGTCAGTACTACTGGCAAATTCGCGTATCGAAGTCCCAGAAGCCGACGAAATTTGCACTCACCCGGCGC  
 GAATCATCAAGGGCACAGTGGTCGAGTCGAGTCGCCACTCACGATATCCGTGCTTACGCGTACGCCTC  
 TCCAAGCCCTTCGAGTTCTACCCGGACAGTACGCGACACTGCAGTTCCAGCCCTGAGCATGCGCGTCCGTA  
 TTCAATGGCAGGTTTGGCAGATGACCAAGAAATGGAGTTCACATACGCAAGGTGCCGGTGGGCGCGTCA  
 CGGAGTATGTTTTCGAACACGTCCGCGAAGGTACAAGCATCAAGTTGAGCGGGCCTTGTGACGGCTTATC  
 TACGTCAAGAAGCACACCCGACCGATGCTGTGTAGGTGGCGGGACCCGACTCGCACCCGTGCTGTGATT  
 GTTCGCGCGCGCTGAAGTCGGGTATGACGAACCCCATCCTCTTTATTTCCGGGTGCGCAGTCAGCAAGA  
 CCTCTACGACGCAGAGCGATTGCACAAACTGCCCGTACCACCCCTCAACTGACCGTACACACGGTGATTG  
 CAACGGGCCGGAATTAATGAGGGTCAGCGAGCCGGCCTAATTACCGATGTGATCGAAAAAGACATCCTTTCCG  
 TGGCTGGGTGGAGGGCCTACCTGTGCGGCGCACCCAGCGATGGTTGAAGCGTTGTCACCCGTCACCAAGCA  
 TCTTGAATATCACCCGAACATATTTATGCCGATGCCTTCTATCCCGTGGGATCTGAATAGTCCCGGCCAT  
 GCACCTCTGTCCATCGAGAATTCATCAGGAAGACATTCAAATGAACGTAACAAATAAGGGCAGCGTCTGTATT  
 TGCGCAGCGAAATGCTCCCTAAATTCCTCATTACCCCATCTGAGGATTGCTTTATGACAGTAAAGTGGATT  
 GAAGCAGTCGCTCTTTCTGACATCCTTGAAGGTGACGTCCTCGGCGTGACTGTGAGGGCAAGGAGCTGGC  
 GCTGTATGAAGTTGAAGGCGAAATCTACGCTACCGACAACCTGTGACGCATGGTTCCGCCCGCATGAGTG  
 ATGGTTATCTCGAGGGTAGAGAAATCGAATGCCCTTGATCAAGGTCCGTTTGACGTTGCACAGGCAAAAG  
 CCCTGTGCGCACCCGTGACACAGAACATCAAACATATCCAGTCAAGATTGAGAACCTGCGCGTAATGATTG  
 ATTTGAGCTAAGATTTTAAACAGGAGGCCACCCGGGCCCTAGAGCGTAATCACCCCATTCATCTTTTTTAA  
 GTGAAAACATGAATTACAATAATAAAATCTGGTAAAGTGAATCTGGTCTGAGCCAAAAGCACCTGATTTCATGG  
 CTGGAAGAACATTTCCAAACATGAACGATAAAACATTTTTGCGCGGAACCTGCTTTCTCACTCATGATAGC  
 CTGATTCCTGCCCGCGGACTATGTTACCGCAAAAATGGGGATTGACGAGGTCATCGTCTCCCGCGCAGAA  
 CGACGGTTCGATTGCTGCTTTTCTGAACGTTTCCCGGCATCGTGGAAGACGCTGGTGAGCGTGAAGCCG  
 GCAATGCCAAAGGTTTTGTTTGCAGCTATCACGGCTGGGGCTTCGGCTCCAACGGTGAACCTGCAGAGCGTT  
 CCATTTGAAAAAGATCTGTACGGCGAGTCGCTCAATAAAAAATGTCTGGGGTTGAAAGAAGTCGCTCGCGTG  
 GAGAGCTCCATGGCTTCATGCTTCCAGTTCGCTTCGACCAGGAGGCCCTCCTCTTATGGACTATCTGGGTGAC  
 GCTGCTTGGTACCTGGAACCTATGTTCAAGCATTCCGGCGGTTTGAACGTGGTCCGCTCCAGGCAAGGTT  
 GTGATCAAGGCCAACTGGAAGGCACCCGCGGAAAACCTTTGTGGGAGATGCATACCACGTGGGTTGGACGCA  
 CGCGCTTCGCTTCGCTCGGGGGAGTCTATCTCTCGTCTGCTGGCAATGCGCGCTACCACCTGAAG  
 GCGCAGGCTTGCAAATGACCTCAAATACGGCAGCGGCATGGGTGTGTTGTTGGACGGATATTCAGGTGTG  
 CATAGCGCAGACTTGGTTCCGGAATTTGATGGCATTCCGAGGCGCAAAGCAGGAAAGGCTGAACAAAGAAAT  
 TGGCGATGTTCCGCGCTCGGATTTATCGCAGCCACCTCAACTGCACCGTTTTCCCGAACACAGCATGCTGAC  
 CTGCTCGGGTGTTTTCAAAGTATGGAACCCGATCGACGCAAACACCACCGAGGCTGACCTACGCCATTGT  
 CGAAAAAGACATGCCTGAGGATCTCAAGCGCCGCTTGGCCGACTCTGTTACAGCAACGTTCCGGGCTGCTG  
 GCTTCTGGGAAAGCAGCAGCAATGACAATGGAACAGCTTCGCAAAACGGCAAGAAATATCAATCAAGAG  
 ATAGTGATCTGCTTTCAAACCTTGTTTTCGGTGAGGACGTATACGGCGACGCGGTCTATCCAGGCGTCTGCG  
 GCAAATCGGCGATCGGCGAGACCAGTTATCGTGGTTTTACCGGGCTTACCAGGCACACGTGACGAGCTCC  
 AACTGGGCTGAGTTGAGCATGCCTCTAGTACTTGGCATACTGAACCTACGAAGACTACTGATCGCTAAACAG  
 ACGAGTCGACCATGATGATCAATATCAAGAAGACAAGCTGGTTTTCCGCCACGACGCCGAAGAGATTTCTTC  
 GTTTCTCAATGCCAGACTCTGCTTTGCAACAAGAAGCCACTACGCTGCTGACCCAGGAAGCGCATTTGTT  
 GGACATTACGGCTTACCCTGCTGTTAGAGCACTGCGTGGGGTCAGAGGTGCAATACAGGTCAATTTCAAG  
 CGAACTGCGCGCAGCTTCAGAGCGTCTTATAAGCTCAATGAAGCCATGAACGTTTACAACGAAAAATTTTCA  
 GCAACTGAAAGTTCCAGTTGAGCATCAACTGGATCCGCAAAACTGGGGCAACAGCCGAAGCTGCGCTTTA  
 CTCGCTTATACCAACGTCCAGGCCGCAATGGACGTAATGACAAAGAGCTACTTCACATCCGCTCCAACG  
 TCATTTGCACCCGGGCACGACTGGCAATCAGGTGATGCTTCTACGCGCCCGGGAAGATAAATGAAA

CGTGGCGAAGGTGGAGTACGAAAATTGGTCCAGCGATTTCGTCGATTACCCAGAGCGCATACTTCAGACGCA  
CAATCTGATGGTCTTTCTGTGA

**CDO gene cluster** (*cumA1* → *cumA2* → *cumA3* → *cumA4*)

GenBank: D37828.1

ATGAGTTCAATAATAAATAAAGAAGTGCAGGAAGCCCTTTGAAATGGGTGAAAACTGGTCTGACGAGGAG  
ATTAAGCGCTCGTTGATGAGGAAAAGGGGTTGCTTGATCCACGATATTTCTCTGATCAGGATTTGATGAGA  
TCGAGCTTGAGAGGGTGTGGCTCGATCTGGCTGCTTGGCAGGAGGGCACATCCCAAGCCGGG  
GATTATCTGACCACCTACTATGGTGAAGACCCAGTAATTGTAGTGAGGCAAGAAAGACCGGAGCATTAAAGTC  
TTTTAAACCAATGTCGGCATCGCGGTATGCGTATTGAGCGATCGGATTTTGGCAACGCAAAGTCATTTACCT  
GCACATTACACGGGTGGCCTATGACACCCGCGGTAATCTGGTCAATGTACCTACGAGAAAGAGGCTTTTT  
TTCGACAAAAAGAGGGTGACTCGCGGTTTCGACAAGGCCGACTGGGGGCCGCTGCAAGCGCGGGTGGATAC  
TTACAAGGGCTGATTTTGGCAACTGGGATACCGAAGCCCTGATTTGAAGACCTATCTGAGCGATGCAAC  
ACCCTATATGGACGTGATGCTCGACTCGGACCGAGGCAGTTACTCAGGTATCACCGGTATGCAAAGACGGT  
AATCCCCTGTAACCTGGAATTCGCCGCGAGCAATCTGTAGCGATATGTACCATGCGGGAACGATGGCGCA  
TCTTTCAGGTGATTTGCCAGCCTCCCGCTGAAATGGATTTGCCAAGTAAAGTTACCGTCAAGTGGGAAT  
CAGTTCGGGGCTAAGTGGGGTGACATGGGACCGGCTGGTCAATGACGATTTGCGACTTCTGCAAGCCAT  
CATGGTCAAGTGTGCGATTACTGACCAAGTCCAGCTGCTGAGCGTGCAAAAGAGCGTCTGGGTA  
AAGTTCCTCGGCTGATCGCATGTTGCTCAGCATATGACCATTTTTCCGACATGCTATTTCTCCTGGCAT  
CAATACAGTCCGTACTTGGCACCCAGTGGCCCTAATGAGATCGAAGTTTGGTCTTTCATCGTAGTGGATGC  
TGATGCACCTGAAGATATCAAGGAAGAATATCGTCGAAAAACATCTTCACCTTCAATCAAGGGGGAACCTAC  
GAGCAGGACGATGGCGAAAAGTGGGTGGAGGTTTCAGCGGGGATTGCGCGGCTACAAGGCTAGAAGTAGAC  
CTCTTTGTGCCAGATGGGGGGGGTGTGCCAAACAAGAACAACCCGGAGTTTCTGGAAAGACCAAGCTAC  
GTTTATAGCGAAGAAGCTGCGCAGGGTCTACCACCCTGGAGCCGATGATGTCGAGCCGAGTTGGGA  
CACGCTAAAGTCTTGAAGCAGATAAAGTACCGAAAAAGCAATCAGTTCATCGGGTTCTACCGTGGTAGAC  
AAGGGTTTAGCCTGTTTTTGGTTGCTGGAAGTGCCTAAGTGAATTGATTAACCTGGGTAAACCCCTGGCTTT  
GTCGGGGTATTTACTCGGGTGCATTCAAAATGTACAGCTGTGCGTTTTGGTGATAATCGTCATGCTATGGA  
TTTGCATTTGCGATGAGCCGAGTGCAGGTCGCCAACATATATACAGAAACTAATTATGACATCCCGTGAAT  
TGACAAAACCCATCGAGTGGCCAGAAATGCCCGTCAAGTCTTGAATTTGCAAAATGCCGTTGAGCAATTTACTA  
TCGCGAAGCACAGTTGCTTGATTATCAAACTATGAGGCCTGGCTGGCTTTACTGACCCAAGACATCCAATAT  
TGGATGCCAATTCGTAATACTATACATCCCGGAATAAGGCGATGGAGTACGTGCCCCCGGGCGTAATGC  
CCATTTTGAAGAGCGTATGAGAGCATGCGTGCAGCATTTCGGCGAGGGTTTCGGGGCTTAACTGGACTG  
AAGATCCACCGTGCAGCCGCGCACATTTGAAGCAACGTTATCGTCCGCAAACTGAGAGTGTGGTACTT  
TGGAAGTTAGTCTGCGTTCCTTTGTTACCGTAATCGATTGGAGCGTATGACGGACATCTATGTCGGTGAGC  
GTGAGATATTTGCTCCGTGAAGTGACGGGCTGGGATTCAAAATGGCAAGCGAACGATCTTGTCTGACC  
AGAGCAGGATTACAGCGAATAATCTCAGCCAGTTTTTCTAACTAGGGAATGCTGGCCACTTACCCATACCCA  
GCTTATCATGAGAGCGGCTGAAAATGAAGGAGGCTACCCGATAGCTACGCAAACTAATCGCGCTCGCC  
CTTTCTGATCGCGATTCGGTATCTTTACTTGGCCAATCTCCTCGGGACTTTGCATTTACGACGCTCGGGCT  
GTTCCGCGATGATGATTCGGGTGTGGATTTGCAGGTCGCGCTCCGGTATTCACCTGCTGCAGGATGCCT  
GGGCCGTAGTCGGGCTGCAGCTGGGGGCGACTGGGCTGGTTGCGTTGTGGGGCGCACGTGAGCCGTGC  
GCTTCATGCGGTTGTCCCGTGGTATCGTCAAGGAAAGTGTGCGACGGTATCTGGGACTTGTACAGCATC  
GTTTGGAGTCACGAAGCCATGTGGTTCGGGCTCCTGACGTTCCCATCCACGTTGGTGTGGATCGTCTGGGG  
GTTACAGGTATGGCGCGTGTGCTGCTCGCGGTCACTGCGCTTAAACCGTCCCAACCTCCTGAATCTGTGGC  
CTGAATGAACCTGTCAATTTCCATGCGCGTGCAGGTTATCCGGCCGCTTGCGCCAAGCTGGCTGCCGGAT  
TTAATGAGTAGTTAAAGTTAGCTATAGAACTCTGAAAAAGGCTTGACCTCATGAGATATCCAGTCTGCAGTC  
CGCGTGGTACTGGCGTGCATTTCCGAGTGCCTACTTTTTAGACCAACTCTATAAAGAGACAAAAAAGA  
ATGACTTTTTCCAAGTTTGAAGTATCTGATGTGCCCGTCCGGTGACGCGCTTGCAGGTTGAAAGTAAGGGC  
GAAGCCGTCGCGATTTTCAACGTCGATGGAGAGTTGTTGCAACACAGAGACCGTTGCACCTAGGTAAGT  
GTCTTGTCCGAAGGCGGTACCTAGAGGGTGACATTGTGCAATGCTCGTGCACATGGGTAGGTTCTGTG  
TCCGACGGGCAAGGTAAGCAGCACCGCCCTGTGAGCCGCTGAAGATATATCCGATTGCAATAGATGGC  
AGCGATGTGTTCTGACTTTGATGCCGGTATCTAGCGCCATGATTAATCAATCGTCATTATGGTCTGG

CTTGGCTGGCGCAACTGCCACTCGCTATCTTCGCGCCCAAGGATATCAGGGAAAGATCCATCTGGTCGGGG  
 AGGAGTTGCATGTGGCTTACGATCGCCCTCCTTATCCAAGGACACCCTGTGAGGAAAAGTGGTCGAACCAC  
 CCGCAATCCTGGATCCTTGTGGTATGCATCGGCCGATATAGATCTCCATTTAGGTGTACCGGTGACCGGTA  
 TTGATGTGGTAAACCACCAGGTACTTTTTGAATCCGGTGACATTCTAGCCTACGACCGACTGCTATTAGCCAC  
 CGGCGCTCGCGCGCGCGGTATGGCTATTACGGGAAGCGAGTTGGCCGGCATTACACCTTTGCGTGAACCGC  
 CGCGACAGCCAGGCGCTGAGGCGAGCGCTTGAGCCGGGCCAGTCTCTGGTAATTGTGGCGGTGGCCCTGA  
 TCGGTTGCGAAGTGGCGACCACTGCTATTAATGCGCGTGCCACGCTACTGTTCTGGAGCCGGGACGAA  
 CTGCTGTTGCGAGTGCTAGGCCGATCAACCGGGCCCTGGTGTGCGCAACGAGTTGGAGCGTTTTGGGTGCC  
 GGGTTGAACTGAACGCACAGGCAGCGCATTTCGAGGGCGAGGGACACGTGCATGCCGTGTTTTGTGCCGAT  
 GGACGTCGGATAGCAGCTGGCACAGTTTTGGTGAGCATCGGTGCAGAACCAGCCGACGAACTGGCACGTG  
 CGGCCGCTATCGCATGTGAGCGCGCGCTGGTAGTTGACGCTACGGGTGCAAGCTCATGCTCAGTAACTC  
 CGCGCAGGTGACGTAGCACCCCTGGCCCTGAGGTCCGGTGAAGCTCGCTGGAGCCCTACCTGAAACA  
 GCCACATGCAGGCTGAAACTGCCGCGCGGCCATGTTAGGCAAGTCTATCCCGGCTCTTACAGTGCCAACC  
 TCTTGGACGGAGATTGCAGGGCATCGGATACAGATGGTTGGCGACATCGAAGGCCCGGAGAAAGTTGTCTT  
 GCGCGGTAACGTCGAGAATGGTCAGCCGCTGGTGCAGTTCAGGGTTCTTGATGGTCGCGTTGAAGCCGCAA  
 CGGCTATCAATGCCCGGAAGATTTCCCGTTGCAACCCGATTGGTGGCTGACCACATTCCTGTATCGGCCA  
 CAAAATTGCAGGACGCTAGCTCTAACTTGCGGGATTTTATGAAAGCTAAAGCTGAGCGGATGCGAGTGA

**BDO gene cluster** (*bedC1* → *bedC2* → *bedB* → *bedA*)

GenBank: AF148496.1

ATGAATCAGACAGAAAACACCTATTAGAGTGCGCAAAAATTGGAAGACCAGTGAGATAGAAACGCTCTTTG  
 ATGAGCAAGCTGGACGTATCGATCCGCGCATTATACCGATGAGGATCTGTACCAACTCGAACTAGAACGTG  
 TATTTGCTCGGTCAATGGCTCCTATTGGGGCATGAAACTCACATTCGTAACACAGGTGATTATTTACAGACCTA  
 TATGGGTGAAGATCCTGCTGGTGCCTGAAGACAGAAGGATGCCAGTATCGCTGTGTTCCGAAACAGTGGCC  
 CCATCTGGTATGCGTATCTCGTTCGGATGCTGGAACCGGAAGGCAATTTCTGTAGTTACCATGGGTG  
 GGCTTACGATACTGCTGGCAATCTTATTAATGTGCCTTACGAGGCCGAATCCTTCGCTGCTTAGACAAAGAA  
 GGAATTGGAGTCCACTGAAGGCTCGAGTGGAAACCTACAAAGTCTGATTTTTGCCAACTGGGATGAAAACGC  
 CATAGACCTTGATACATATCTCGCGCAGGCGAAGTTCTACATGGACCACATGCTCGACCGTACTGAGGCAGG  
 CACTGAGGTTGATCCAGGTATACAGAAGTGGGTTATCCCTGTAAGTGGAAATTCGCGACGAGAGTGGTTTTG  
 TAGTGACATGTACCATGCGCGGACTACAGCACATTTATCTGGAATCATTGCTGGTCTGCCAGAAGACTTGA  
 GTTGGCTGATCTTGACCCGCCGAAATTTGGCAAGCAGTACCGTGCATCATGGGGTGGGCATGGCAGTGGCT  
 TCTATATTGGCGACCCCAACATGATGCTTGCCATGATGGGGCCGAAGTCCACGACTACTTGACCGAAGGC  
 CCCGCGCGGAAAAGGCGCGCGAGCGCTGGGTAGTATAGAGCGCGCACGAAAATCATGCTTGAGCACA  
 TGACTGTCTTCTACGTGTTCTTCTCCAGGTGCAATACGATCCGAACATGGCATCCACGCGGGCCGA  
 ACGAGGTTGAAGTGTGGGCATTTACAGTCGTGATGCTGATGCTCCAGATGATATTAAGGAAGATTTTCGTC  
 GTCAGACACTACGTACCTTCTGCGCGTGGTGTATTGAGCAGGATGACGGCGAGAAGTGGGTTGAAATC  
 CAGCATATTCTGCGAGGTCATAAGGCAGTAGCCGTCCATTCAATGCTGAGATGAGTATGGGGCAAACCGTT  
 GATAACGATCCAATTTACCTGGTGTATATCTAACAACGTTGATAGCGAAGAAGCTGCTCGTGGACTATATG  
 CACATTTGGCTGAAAATGATGACCTCACCAGACTGGGAAGCATTAAAGGCGACGCGTTAAACCTAACGACAGC  
 TAGCGCCACATCGTGGTCCAGTTAGAAGCCGCAATTTGATTCAACAAAATTGGATGCGGTGACTATCCCA  
 TTTGAAATCTACACGGAATGATTACGATTCCAAAAGGAGTGACCATTGATTGATTGATTGATTGATTGATTGATTG  
 TCTTTCTCGCAAACCTGCACCGGTAGCACTTGAAGTGCAAAATGAAATGAGCAGTTCTACTATTGGGAAGC  
 TAAGCTTCTCAATGATCGCGCTTTGATGAATGGTTGCGCACTACTTGCCAAAAGACATTCACTACTTCATGCT  
 ATCCGCACCACAGAAATCATGCGTGATTACGCGCTTGAATATTCGGGCTTGGCAGACTATGCACATTTTGTG  
 ATGACGCCACAATGATGAAAGGACGTTTGGCGTAAGATTACTTCTGACGTAAGTTGCTGCGAAGTCCGGAATCCTGCT  
 CAAGAACACGACATATTGTGAGCAATGTGATGATCATTCACAGAAAGTAGAAGGAGAATACGAAATCTCCA  
 GTACCTTACGTGTACCGCAATCGCTTGGAAAGACAACCTTGATATCTTGTGAGAGCGTCCGCGACAGAT  
 TGCGTCGTAACAAAGGGTGAAGCTGGATTCGAGATAGTCAATCGGACAACTCCTGATCGACAAAAGCACCATC  
 TAGCCAATAACCTCAGTTTCTTCTTAGGAGATGTCATGACTTGGACATATACTGCGGCAAAGTGAATTTG  
 CCACCTGGTGAATGCAACGCTATGAAGTGGATCAGAACCAGTGTGCTGTGTAACGTCGATGGTGAATTT  
 TTCGAGTTTCAAGATACCTGCACGATGGAGACTGGGCATTATCAGAGGGATACCTTGATGGTGTGCTGCTG

GAATGTACGTTGCATTTTGGGAAGTTCTGTGTGCGAACTGGGAAAGTGAAAGCGTTGCCTGCTTGTAAACCT  
ATCAAAGTATACCCTATCAAGATAGAAGGTGATGAGGTACACGTTGATCTTGACAATGGGGAGCTAAAGTGAT  
GGCTAATCATGTTGCAATCATCGGTAATGGCGTAGCTGGGTTACCCACCGCGCAAGCCCTTCGTGCCGAAG  
GTTATGAAGGACGAATCTCACTGATTGGGGAAGAACAGCATCTCCCTATGACCGACCATCCTTGTCTAAGG  
CCGTCCCTTGATGGAAGCTTTGAGCAGCCACCTAGACTGGCCGAGGCAGATTGGTACAGTGAAGCCAGCATC  
GAGATGCTGACTGGCTCAGAAGTTACTGACCTAGATACACAGAAAAAATGATCAGTTTGAATGATGGCAGC  
ACGATTTCTGCTGATGCTATCGTTATTGCAACTGGAAGTCGAGCACGGATGTTGCTTTACCCCGCAGTCAAT  
TGCCCTGGTGTGCGTACATTACGCACTTATGGTGTGTCAGTTATTACGTGATAGCTGGACACCCAATACTC  
GGTTGCTCATTGTAGGCGGTGGGTTGATTGGTTGTGAGGTAGCAACAACGGCTCGCAAGCTTGGCTTTTCTG  
TCACGATCCTTGAAGCTGGTGTGAACTGTTGGTTCGTGTTCTTGGACGACGTATCGGTGCTTGGCTACGTG  
GTTTGTAACTGAACAGGGTGTGCAGGTAGAATAAAAACTGGAGTTTCAGGTTTTTCAGGCGAAGGTCAGC  
TTGAAAAAGTGATGGTGAATGATGGGCGTAGCTTTATTGCTGATAACGCACATTATCTGCGTAGGGGCAGATC  
CTGCAGATCAACTCGCACGTCAAGCCGGTTTGGAAATGTGATCGTGGTGTGCTTGTGATCATAGAGGTGCGA  
CATCTGCTAAAGGCATATTCGCGGTGCGGGATGTAGCCACTTGGCCACTTCATTAGGCGGAAAGCGCTCG  
CTTGAAACCTATATGAATGCTCAGCGCCAAGCCACAGCAGTTGCCAAAGCCATTCTAGGAAAAGAAGTATCA  
GCACCGCAATTGCCAGTGTGATGGACGGAGATTGCCGGGCATCGAATGCAAAATGGCTGGTGTATCGAAGG  
ACCGGGTGAATATGTTTTGCGCGGAACCTTAGGCATTGGCTCTGCTTTATTGTTCCGCTGCTAGATGGACG  
GATTCAGCGGTTGTAGCGGTTGATGCACCTCGTATTTCGCACTCGCAAAATCGATTAGTGAAGCTCAAGT  
CATAATTGAGCCAGAAAAGCTGGCAGATGTGTCAAATAATATGCGCGATATTGTTCTGCGAATGAAGGGAA  
TCAAAAATGA

### 7.4. Sequence alignments

#### CustalW2 multiple sequence alignment<sup>220</sup>

```

BDO      MN----QTETTP IRVRKNWKTSEIETLFDEQAGRIDPRIYTDDELYQLELERVFARSWL 55
TDO      MN----QTDTPS IRLRRSWNTSEIETALFDEHAGRIDPRIYTDDELYQLELERVFARSWL 55
CDO      MSSIIINKEVQEAPLKWVKNWSDEEIKALVDEEKGLLDPRIYFSDQDLYEIELERVFARSWL 60
NDO      -----MNYNN---KILVSESGLSQKHLIHGDEELFQHELKTI FARNWL 40
          .. : *..* . * .*:*: : * : * : * : *

BDO      LLGHETHIRKPGDYFTTYMGEDPVVVVRQKDASIAVFLNQCRHRGMRI CRSDAGNAKAFT 115
TDO      LLGHETQIRKPGDYITTYMGEDPVVVVRQKDASIAVFLNQCRHRGMRI CRADAGNAKAFT 115
CDO      LLGHEGHIPKAGDYLTTYMGEDPVVVVRQKDRSIKVF LNQCRHRGMRIERSDFGNAKSFT 120
NDO      FLTHDSLIPAPGDYVITAKMGIDEVIVSRQNDGSI RAFLNVCRHRGKTLVSV EAGNAKGFV 100
          : * * : * . * * : * * * * * * * * * * * * * * * * * * * * * * *

BDO      CSYHGWAYDTAGNLINVPYEAESFA-----CLDKKEWSP-LKARVETYKGLIFANWDE 167
TDO      CSYHGWAYDTAGNLVNPYEAESFA-----CLNKKEWSP-LKARVETYKGLIFANWDE 167
CDO      CTYHGWAYDTAGNLVNPYKEEAFCDKKEGDCGFDKADWGP-LQARVDTYKGLIFANWDT 179
NDO      CSYHGWGFGSNGELQSVPF EKDLYG-----ESLNKKLGLKEVARVESFHGFIYCGFDQ 154
          * : * * * . : : * * * * * : : * * . * * * : : * * * : : * *

BDO      NAIDLDTYLGEAKFYMDHMLDRTEAGTEVIPGIQKWI PCNWKFAAEQFCSDMYHAGTTA 227
TDO      NAVDLDTYLGEAKFYMDHMLDRTEAGTEAIPGVQKWI PCNWKFAAEQFCSDMYHAGTTS 227
CDO      EAPDLKTYLSDATPYMDVMLDRTEAVTQVITGMQKTVIP CNWKFAAEQFCSDMYHAGTMA 239
NDO      EAPPLMDYLGDAAWYLEPMPKHS GG-LELVGP PGKVIKANWKAPAENFVGDAYHVW-WT 212
          : * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

BDO      HLSGIIAG---LPEDLELADLAPPKFKQYRASWGGHSGS GFYIGDPNMLLAMMPKVTYSY 284
TDO      HLSGIIAG---LPEDLEMADLAPP TVGKQYRASWGGHSGS GFYVGDPNMLLAIMGPKVTYSY 284
CDO      HLSGVLSS---LPPEMDSLQVKLPSSGNQFRAK WGGHGTGWFNDDFALLQAIMGPKVVYD 296
NDO      HASSLRSGESIFSSLAGNAALPPEGAGLQMTSKYSGSGMGLW DYGSGVHSADLVP ELMAF 272
          * * . : . : . : : * * : : * . : . : * : * : :

BDO      LTEGPAEAKAERLGSIERGTKIMLEHMTVPPTCSFLPGVNT IRTWHPRGPNVEVWVAF 344
TDO      WTEGPAEAKAERLGSVERGSKLMVEHMTVPPTCSFLPGINTVRTWHPRGPNVEVWVAF 344
CDO      WTKGPAERAKERLGKVLPA DRMVAQHMTIFPTCSFLPGINTVRTWHPRGPNIEVWVSI 356
NDO      G--GAKQERLNKEIGDVRARIYRSHLNCTVFPNNSMLTCSGVFKVWNP IDANTTEVWTYA 330
          * . * : : * : : : * * * * * * . . . . : * * * * *

BDO      VVDADAPDDIKEEFRRQLR TFSAGGVFEQDDGENWVEIQHILRGHKARSRPFNAEMSMG 404
TDO      VVDADAPDDIKEEFRRQLR TFSAGGVFEQDDGENWVEIQHILRGHKARSRPFNAEMSMG 404
CDO      VVDADAPEDIKEEYRRKNIFTFNQG GTYEQDDGENWVEVQRGLRGYKARSRPLCAQMAG 416
NDO      IVEKDMPEDLKRR LADSVQRTFGPAGFWE SDDNDNMETASQNGKKYQSRDSDL LSNLFGF 390
          : * : * * * * * . . * * . * * * * * * * . : : : * . : : : .

BDO      QTVNDNDIYPGRISNNVYSEEAARGLYAHWLKMMTSPDWEALKATR----- 450
TDO      QTVNDNDVYPGRISNNVYSEEAARGLYAHWLRMMTSPDWDALKATR----- 450
CDO      VPKNNPEFP GKTS-YVYSEEAARGFYHHWRMMS EP SWDTLKS----- 459
NDO      EDVYGDAVYPGVVGKSAIGETS YRGFYRAYQAHVSSSNWAEFHASSTWHTELTKTTDR 449
          . : . * * . . . * : * * * * : : : * * :
    
```

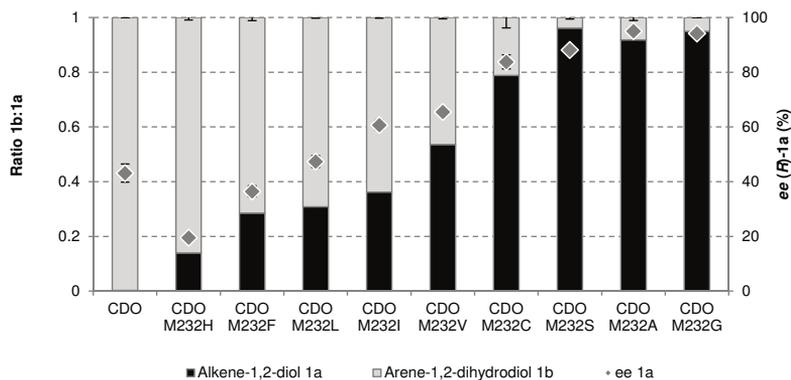
**Figure S30:** Multiple sequence alignment of BDO (*P. putida* ML2), CDO (*P. fluorescens* IP01), NDO (*Pseudomonas* sp. NCIB 9816-4) and TDO (*P. putida* F1) α-subunit amino acid sequences performed with ClustalW2. Amino acid residues that correspond to position 232 in CDO are highlighted in yellow.





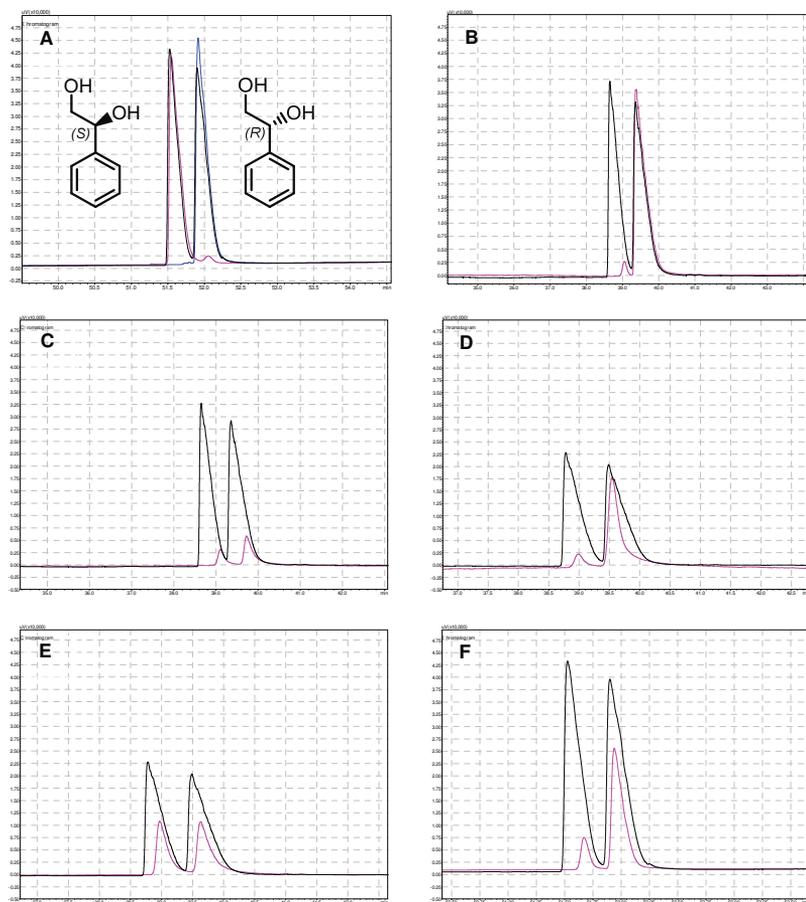


## 7.5. Stereoselectivities

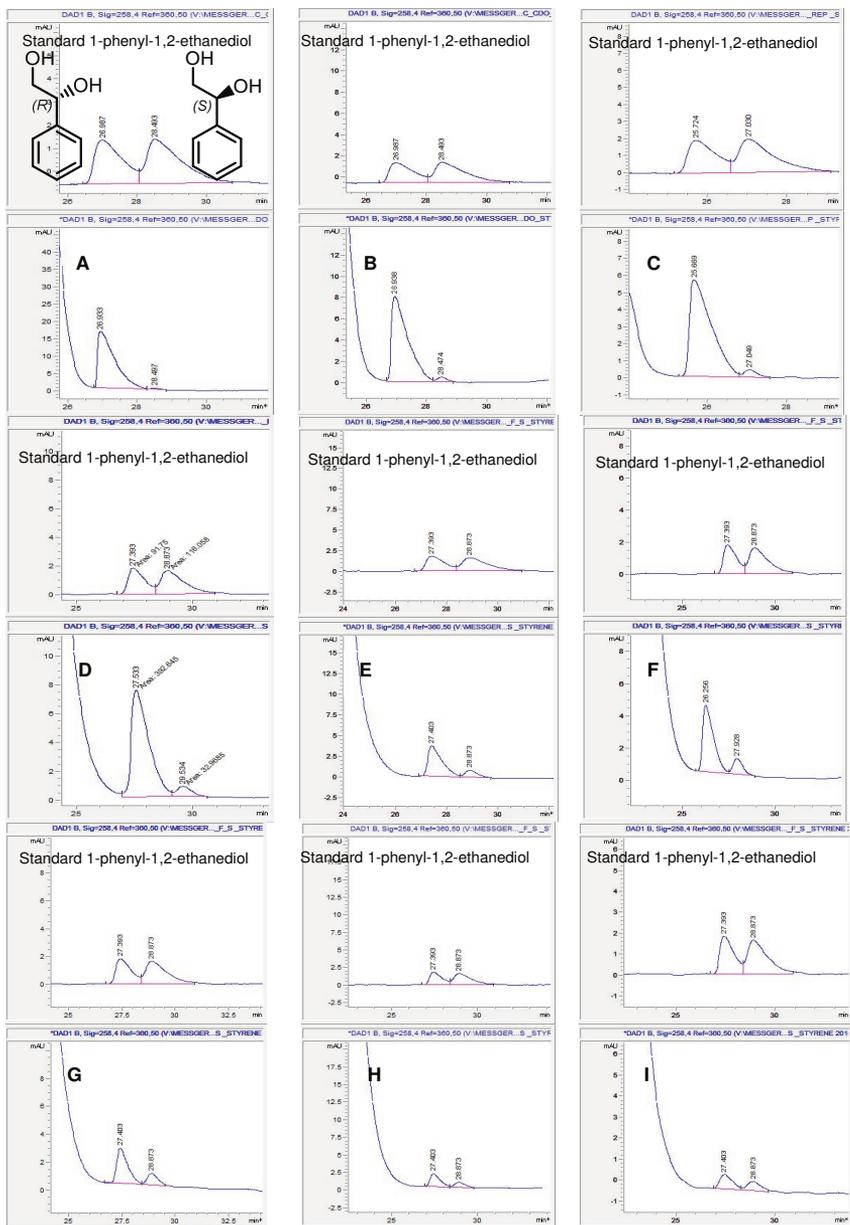
1-Phenyl-1,2-ethanediol (**1a**)

**Figure S31:** Product distribution (bars) for the RO-catalyzed conversion of styrene **1** yielding alkene-1,2-diol **1a** (black) and arene-1,2-dihydrodiol **1b** (grey) as well as stereoselectivities for **1a** (squares) with the variants of the NDT mutant library as well as CDO wild type and CDO M232A.

## 1-Phenyl-1,2-ethanediol (1a)

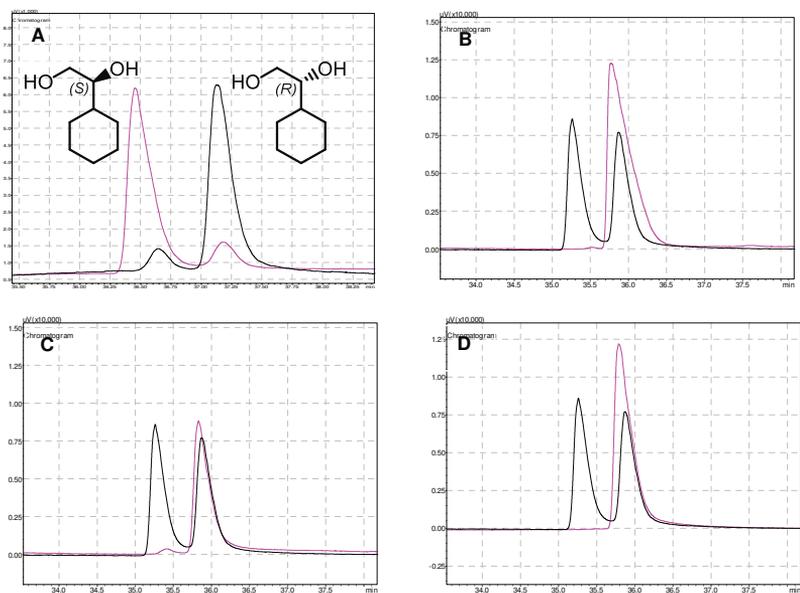


**Figure S32:** Chiral GC-FID analysis of 1-phenyl-1,2-ethanediol **1a** enantiomers of the commercially available standard (**A**) or from RO-catalyzed conversion of styrene **1** (**B-F**). **A:** *rac*-**1a** (black, Fluka), (*R*)-**1a** (blue, Sigma-Aldrich, 99% optical purity), (*S*)-**1a** (pink, Sigma-Aldrich, 99% optical purity). **B:** CDO M232A, **C:** CDO, **D:** NDO, **E:** BDO, **F:** BDO M220A (**B-F** shown in pink compared to *rac*-**1a** in black).

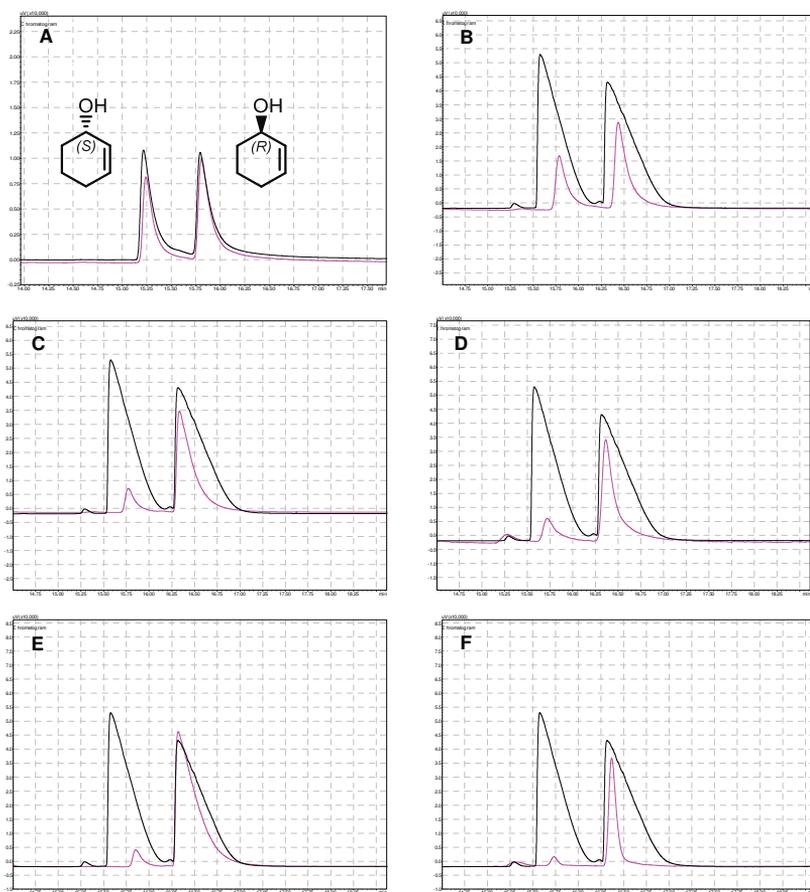


**Figure S33:** Chiral HPLC analysis of 1-phenyl-1,2-ethanediol **1a** enantiomers of the commercially available standard *rac-1a* (Fluka, upper panel) or from RO-catalyzed conversion of styrene **1** (A-I, lower panel). **A:** CDO M232A, **B:** CDO M232G, **C:** CDO M232S, **D:** CDO M232C, **E:** CDO M232V, **F:** CDO M232I, **G:** CDO M232L, **H:** CDO M232F, **I:** CDO M232H.

1-Cyclohexylethane-1,2-diol (**2a**)



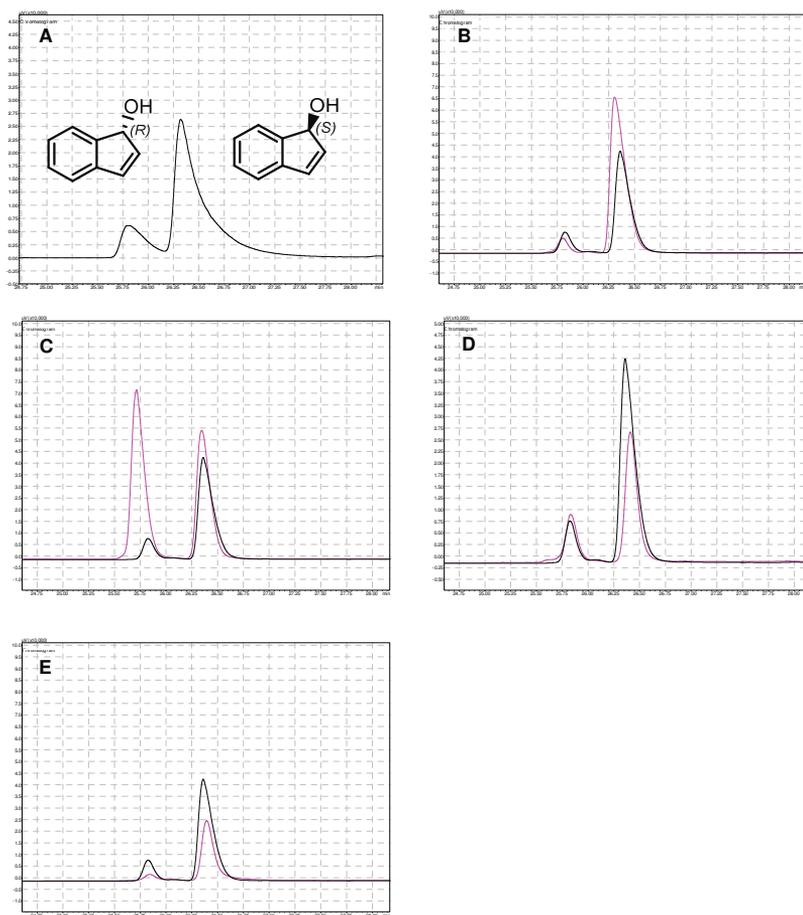
**Figure S34:** Chiral GC-FID analysis of 1-cyclohexylethane-1,2-diol **2a** enantiomers of the synthesized standard (**A**) or from RO-catalyzed conversion of vinylcyclohexane **2** (**B-D**). **A:** (*R*)-**2a** (black), (*S*)-**2a** (pink). **B:** CDO M232A, **C:** CDO, **D:** NDO (B-D shown in pink compared to *rac*-**2a** in black).

2-Cyclohexenol (**3a**)


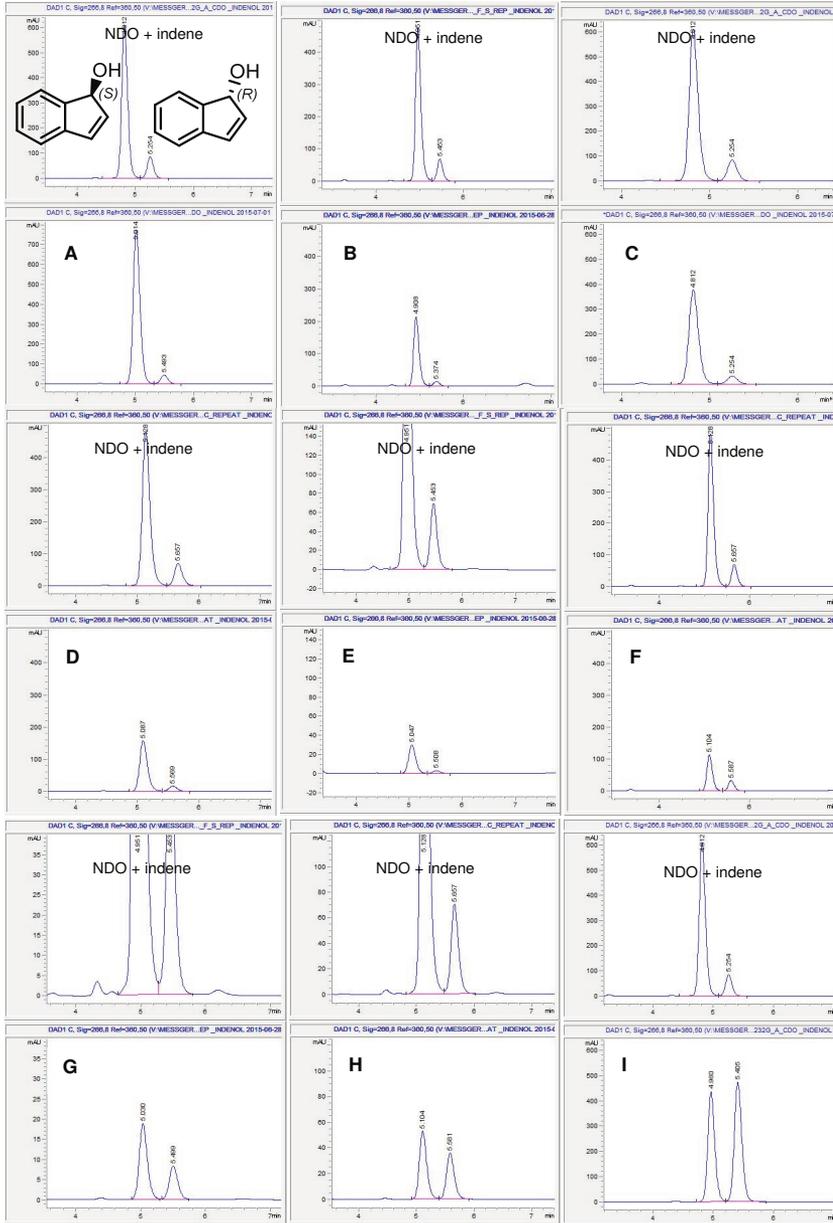
**Figure S35:** Chiral GC-FID analysis of 2-cyclohexenol **3a** enantiomers of the synthesized standard (**A**) or from RO-catalyzed conversion of cyclohexene **3** (**B-F**). **A:** *rac*-**3a** (black, Fluka), synthesized *R*-**3a** (pink, *ee* = 13%), **B:** CDO M232A, **C:** CDO, **D:** NDO, **E:** BDO, **F:** BDO M220A (B-F shown in pink compared to *rac*-**3a** in black).

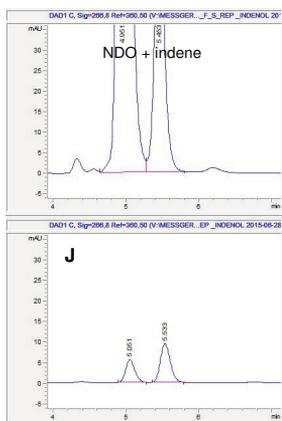
**1-Indenol (4a)**

1-Indenol enantiomers were identified based on comparison with data known from literature for the NDO-catalyzed conversion of indene **4** yielding (S)-indenol.<sup>92</sup>



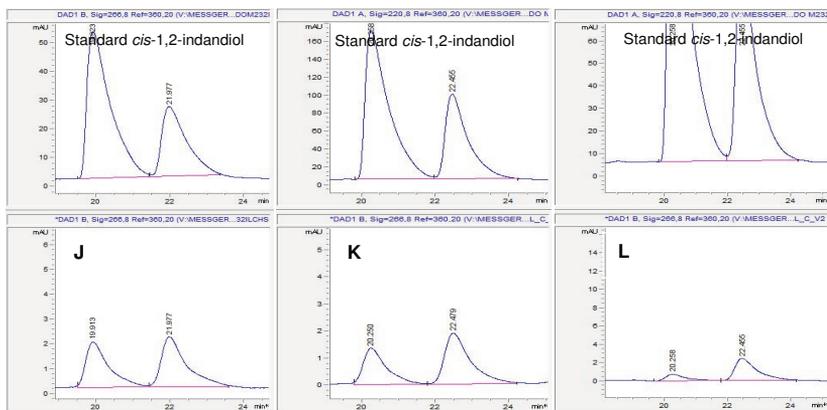
**Figure S36:** Chiral GC-FID analysis of 1-indenol **4a** enantiomers from RO-catalyzed conversion of indene **4** (A-F). **A:** NDO (black), **B:** CDO M232A, **C:** CDO, **D:** BDO, **E:** BDO M220A (**B-E** shown in pink compared to NDO in black).





**Figure S37:** Chiral HPLC analysis of 1-indenol **4a** enantiomers from RO-catalyzed conversion of indene **4** (**A-J**). 1-Indenol **4a** formation of NDO (upper panel), **A**: CDO M232A, **B**: CDO M232S, **C**: CDO M232G, **D**: CDO M232C, **E**: CDO M232V, **F**: CDO M232I, **G**: CDO M232F, **H**: CDO M232L, **I**: CDO, **J**: CDO M232H (lower panel).

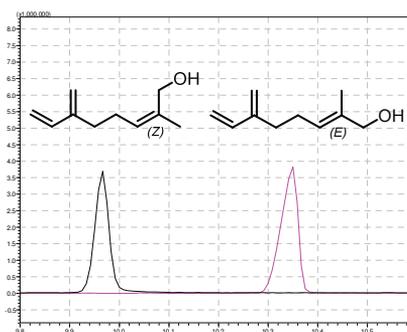




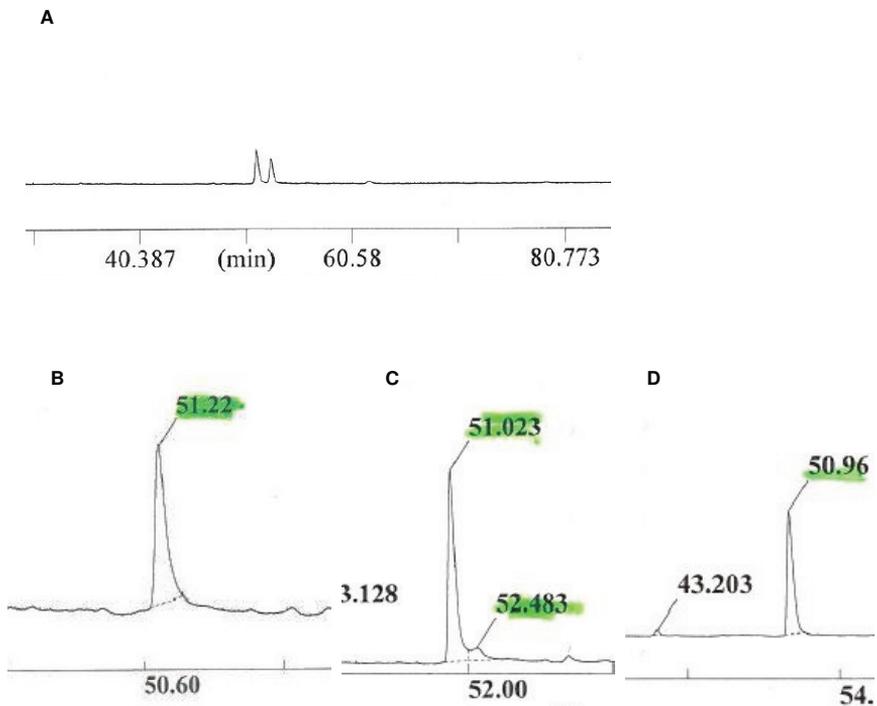
**Figure S38:** Chiral HPLC analysis of *cis*-1,2-indandiol **4b** enantiomers synthesized with AD-mix  $\beta$  (*ee* = 46% for (1*R*,2*S*)-**4b**; upper panel) and of the RO-catalyzed conversion of indene **4** (**A-L**, lower panel). **A:** CDO M232A, **B:** CDO, **C:** NDO, **D:** BDO, **E:** BDO M220A, **F:** CDO M232S, **G:** CDO M232G, **H:** CDO M232C, **I:** CDO M232V, **J:** CDO M232F, **K:** CDO M232I, **L:** CDO M232L.

### 2-Methyl-6-methyleneocta-2,7-dien-1-ol (**5a**)

The product 2-methyl-6-methyleneocta-2,7-dien-1-ol (**5a**) of the CDO M232A-catalyzed biotransformation of myrcene **5** could be identified as (*Z*)-isomer by comparison of its <sup>1</sup>H NMR spectrum with the synthesized standard (*E*)-2-methyl-6-methyleneocta-2,7-dien-1-ol and the different retention times of both compounds in GC-analysis albeit their MS fragmentation patterns are identical.

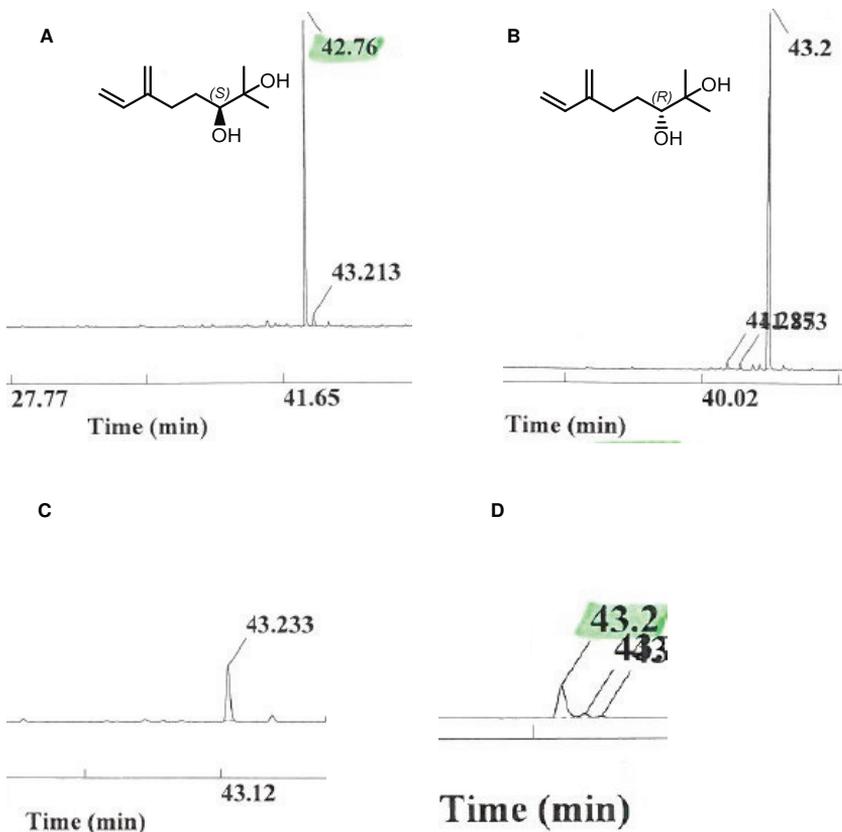


**Figure S39:** GC-MS analysis of (*Z*)- and (*E*)-2-methyl-6-methyleneocta-2,7-dien-1-ol **5a** on a achiral DB-5 column. RO-catalyzed biotransformation of myrcene **5** to yield (*Z*)-**5a** is shown in black, (*E*)-**5a** obtained from synthesis with  $\text{SeO}_2$  and myrcene is shown in pink.

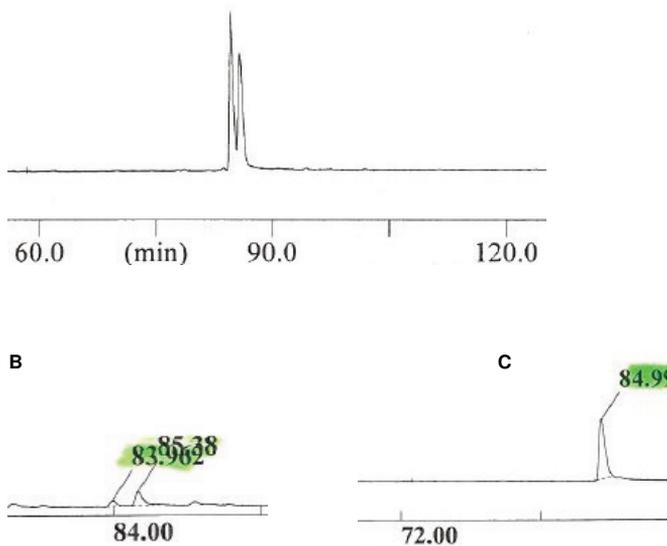
**3,10-dihydro myrcene (5b)**

**Figure S40:** Chiral GC-FID analysis of 3,10-dihydroxy myrcene **5b** enantiomers synthesized with potassium permanganate (**A**) or from RO-catalyzed conversion of myrcene **5** (**B-D**). **B:** CDO M232A, **C:** CDO, **D:** NDO. As dihydroxylation with potassium permanganate was not stereoselective, absolute configurations could not be determined.

2-Methyl-6-methyleneoct-7-ene-2,3-diol (**5c**)



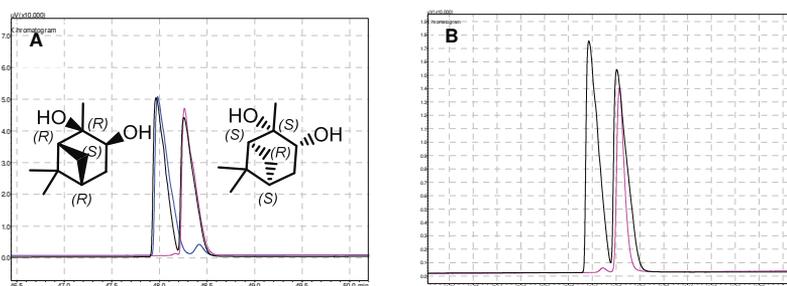
**Figure S41:** Chiral GC-FID analysis of 2-methyl-6-methyleneoct-7-ene-2,3-diol **5c** enantiomers synthesized with AD-mix  $\alpha$  (A) and  $\beta$  (B) or from RO-catalyzed conversion of myrcene **5** (C+D). C: CDO M232A, D: CDO.

1,2-dihydro myrcene (**5d**)

**Figure S42:** Chiral GC-FID analysis of 1,2-dihydro myrcene **5d** enantiomers synthesized with potassium permanganate (**A**) or from RO-catalyzed conversion of myrcene **5** (**B+C**). **B**: CDO, **C**: NDO. As dihydroxylation with potassium permanganate was not stereoselective, absolute configurations could not be determined.

**Pinanediol (6a)**

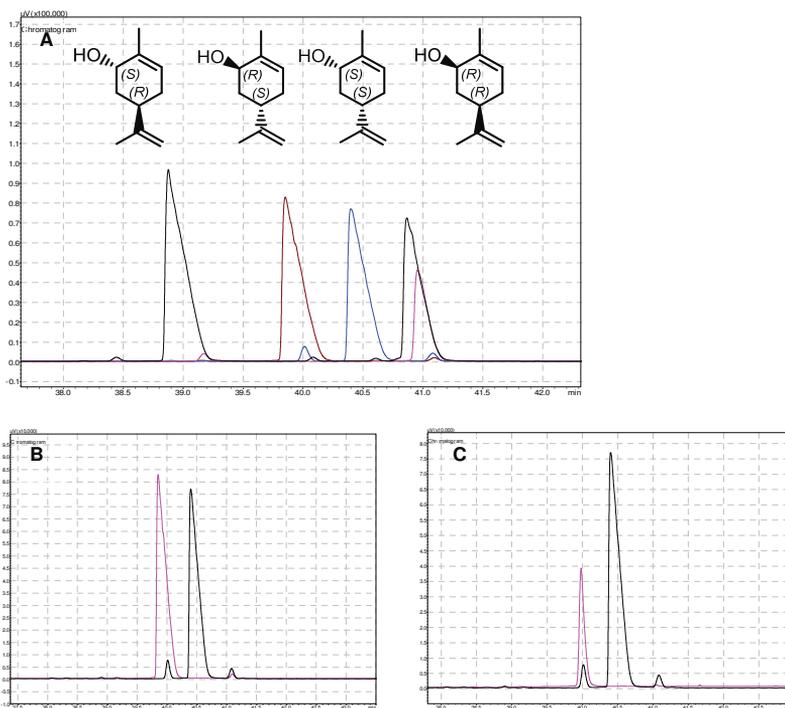
As potential products during the CDO M232A-catalyzed biotransformation of (+)- $\alpha$ -pinene **6**, (1*S*,2*S*,3*R*,5*S*)-(+)- and (1*R*,2*S*,3*R*,5*S*)-(+)-pinanediol **6a** could be formed. However, racemic synthesis of (1*S*,2*S*,3*R*,5*S*)-(+)- and (1*R*,2*S*,3*R*,5*S*)-(+)-pinanediol diastereomers was not possible by osmylation and only the (1*S*,2*S*,3*R*,5*S*)-(+)-isomer could be obtained. Thus, a chiral GC method was developed using the commercially available (1*S*,2*S*,3*R*,5*S*)-(+)- and (1*R*,2*R*,3*S*,5*R*)-(-)-pinanediol enantiomers. As this method allows the separation of enantiomers, diastereomer separation should also be possible. As during chiral GC-analysis, only the (1*S*,2*S*,3*R*,5*S*)-(+)-isomer could be detected in biotransformations using CDO M232A, in addition, a  $^1\text{H}$  NMR and a  $^1\text{H}$   $^1\text{H}$  COSY analysis of the product was performed. According to coupling constants at 3.99 ppm (1 H, dd,  $J_1 = 9.36$  Hz,  $J_2 = 5.20$  Hz), no detectable amounts of (1*R*,2*S*,3*R*,5*S*)-(+)-pinanediol were present ( $de > 95\%$ ).



**Figure S43:** Chiral GC-FID analysis of pinanediol **6a** enantiomers of the commercially available standard **(A)** or from CDO M232A-catalyzed conversion of (+)- $\alpha$ -pinene **6** **(B)**. **A:** *rac*-**6a** (black), (1*S*,2*S*,3*R*,5*S*)-(+)-**6a** (pink, Sigma-Aldrich), (1*R*,2*R*,3*S*,5*R*)-(-)-**6a** (blue, Sigma-Aldrich). **B:** CDO M232A (pink) compared to *rac*-**6a** (black).

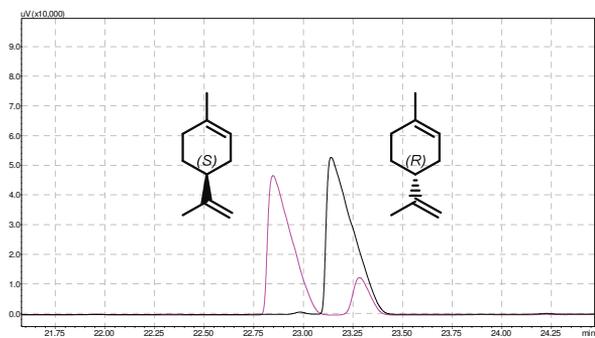
**Carveol (7b)**

(*R*)-(+)-limonene (*R*)-**7** (Sigma-Aldrich) was available with an *ee* of 99%. (*1R,5R*)-(-)-**7b** and (*1S,5R*)-(-)-**7b** that were formed in traces as a result of (*S*)-(-)-**7** hydroxylation were not considered for determination of *de* values.



**Figure S44:** Chiral GC-FID analysis of carveol **7b** isomers of the synthesized standards (**A**) or from RO-catalyzed conversion of (*R*)-limonene (*R*)-**7** (**B+C**). **A:** Black: racemic mixture of (-)-carveol isomers (Sigma-Aldrich), (*1R,5R*)-(-)-**7b** (pink), (*1S,5S*)-(+)-**7b** (blue), (*1R,5S*)-(+)-**7b** (brown) from biotransformations using CDO M232A and (*R*)-**7b**. **B:** CDO M232A, **C:** CDO (**B+C** shown in pink compared to (*1S,5S*)-(+)-**7b** in black).

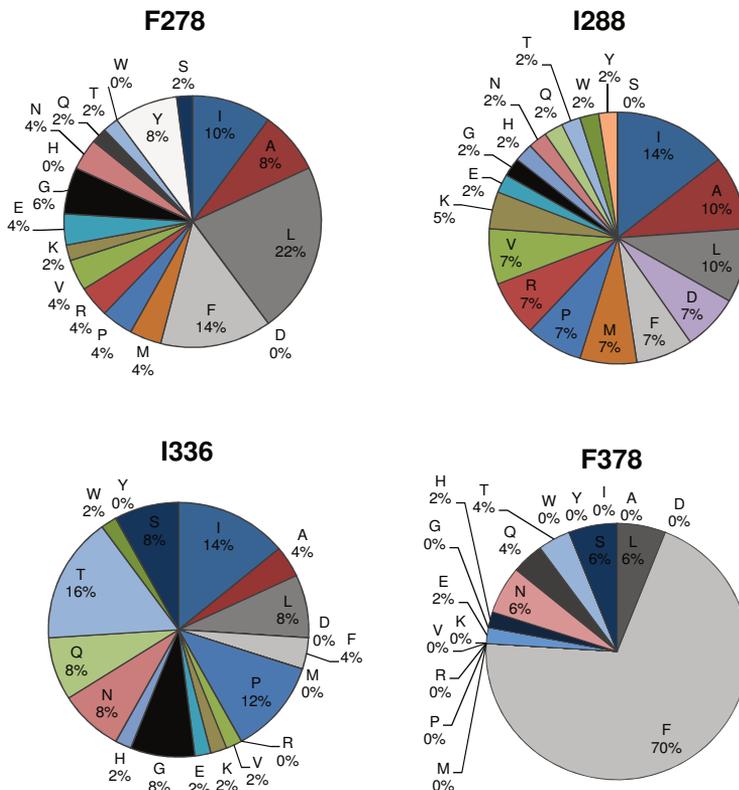
## Limonene (7)



**Figure S45:** GC-FID analysis of the chiral separation of limonene enantiomers. Black: (*R*)-(+)-limonene (Sigma-Aldrich), pink: (*S*)-(-)-limonene (Alfa Aesar). *ee* (*S*)-(-)-limonene = 68%, *ee* (*R*)-(+)-limonene = 99%

### 7.6. Conservation level of positions F278, I288, I336 and F378 in CDO

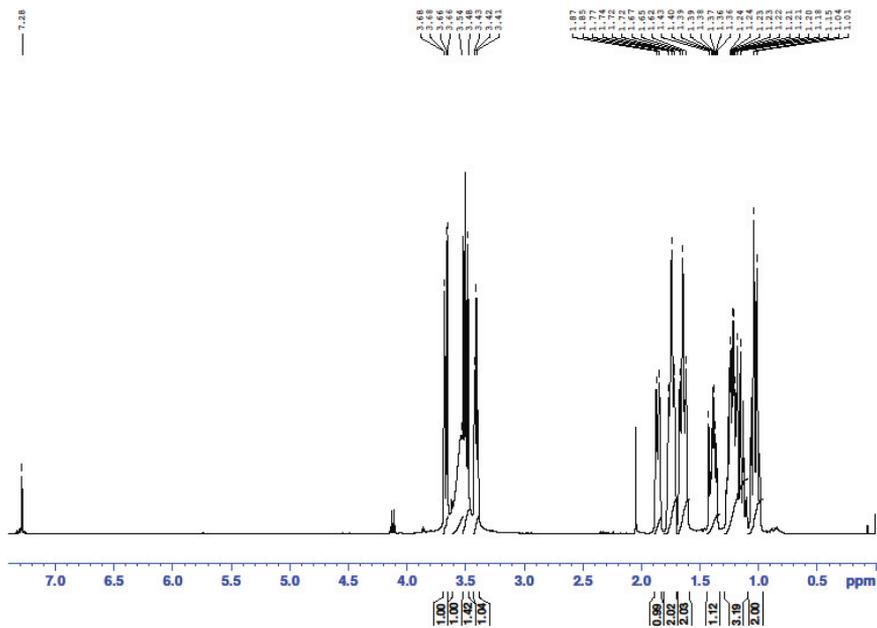
The conservation level of residues F278, I288, I336 and F378 (numbering refers to CDO) chosen for site-saturation mutagenesis and the amino acid distribution at the respective position was determined by HotSpot Wizard.<sup>184</sup>



**Figure S46:** Conservation level and amino acid distributions at positions F278, I288, I336 and F378 (numbering refers to CDO) as determined by HotSpot Wizard.<sup>184</sup>

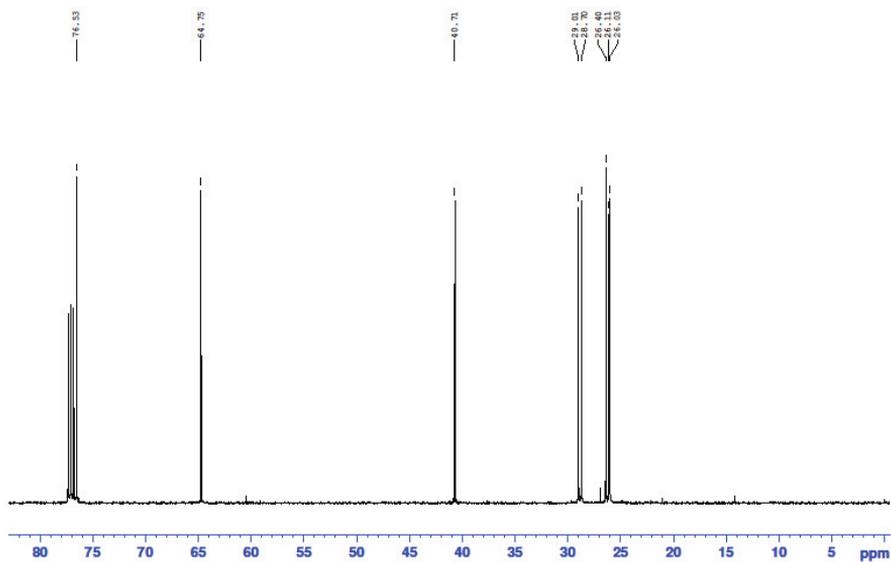
## 7.7. NMR spectra

<sup>1</sup>H NMR 1-cyclohexylethane-1,2-diol (**2a**, CDO M232A)



**Figure S47:** <sup>1</sup>H NMR of 1-cyclohexylethane-1,2-diol in CDCl<sub>3</sub> isolated from the CDO M232A-catalyzed biotransformation of vinylcyclohexane.

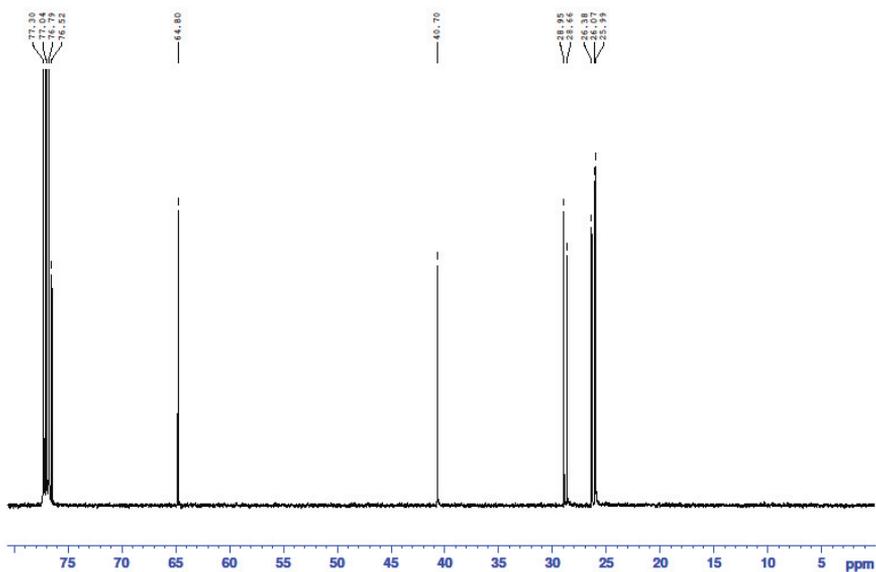
$^{13}\text{C}$  NMR 1-cyclohexylethane-1,2-diol (**2a**, CDO M232A)



**Figure S48:**  $^{13}\text{C}$  NMR of 1-cyclohexylethane-1,2-diol in  $\text{CDCl}_3$  isolated from the CDO M232A-catalyzed biotransformation of vinylcyclohexane.

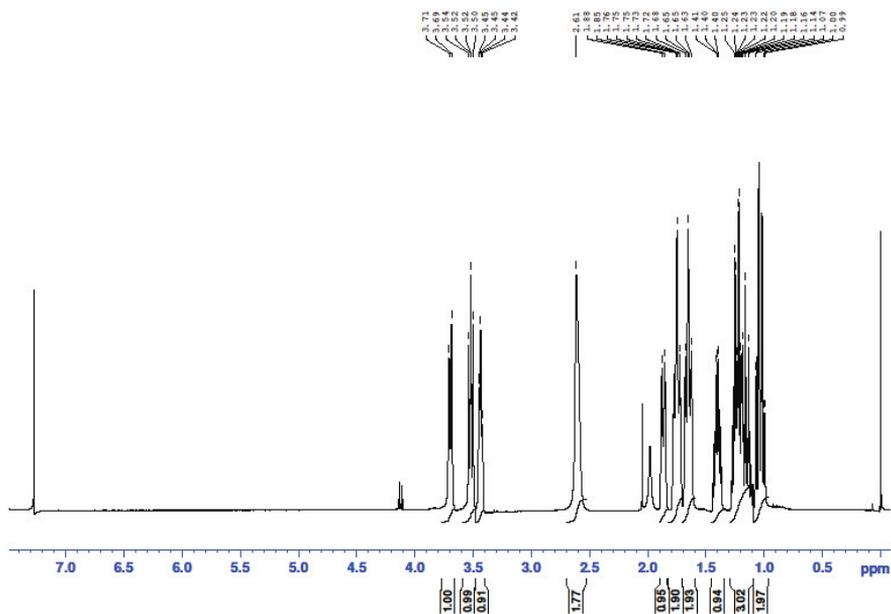


$^{13}\text{C}$  NMR 1-cyclohexylethane-1,2-diol (**2a**, AD-mix  $\alpha$ )

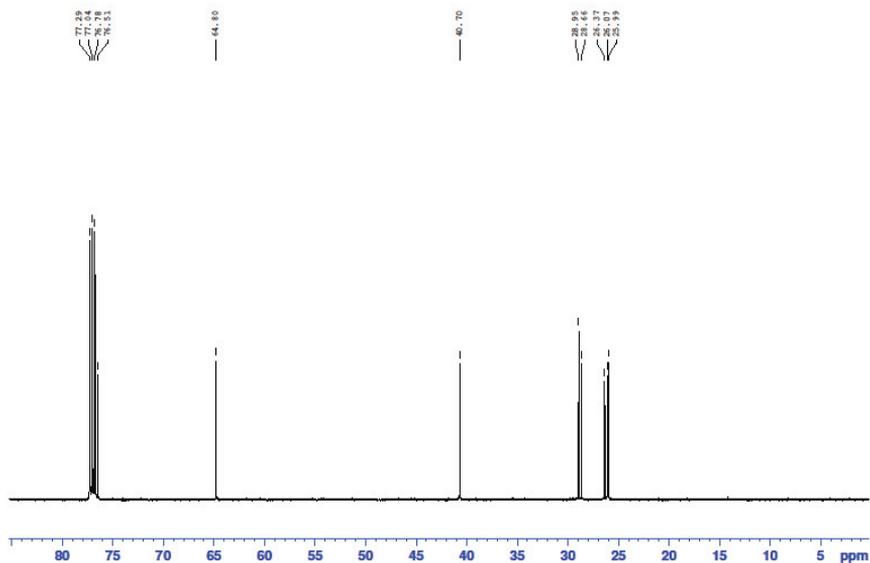


**Figure S50:**  $^{13}\text{C}$  NMR of 1-cyclohexylethane-1,2-diol in  $\text{CDCl}_3$  isolated from the AD of vinylcyclohexane using AD-mix  $\alpha$ .

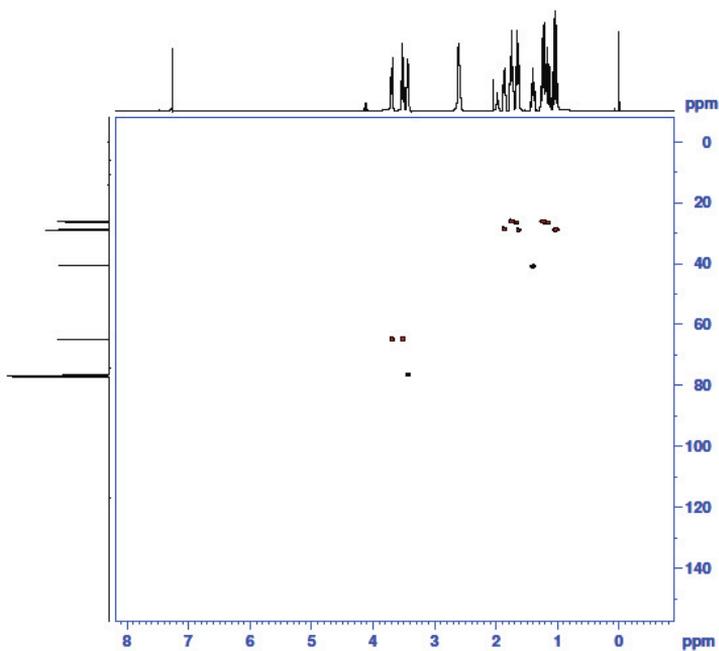
$^1\text{H}$  NMR 1-cyclohexylethane-1,2-diol (**2a**, AD-mix  $\beta$ )



**Figure S51:**  $^1\text{H}$  NMR of 1-cyclohexylethane-1,2-diol in  $\text{CDCl}_3$  isolated from the AD of vinylcyclohexane using AD-mix  $\beta$ .

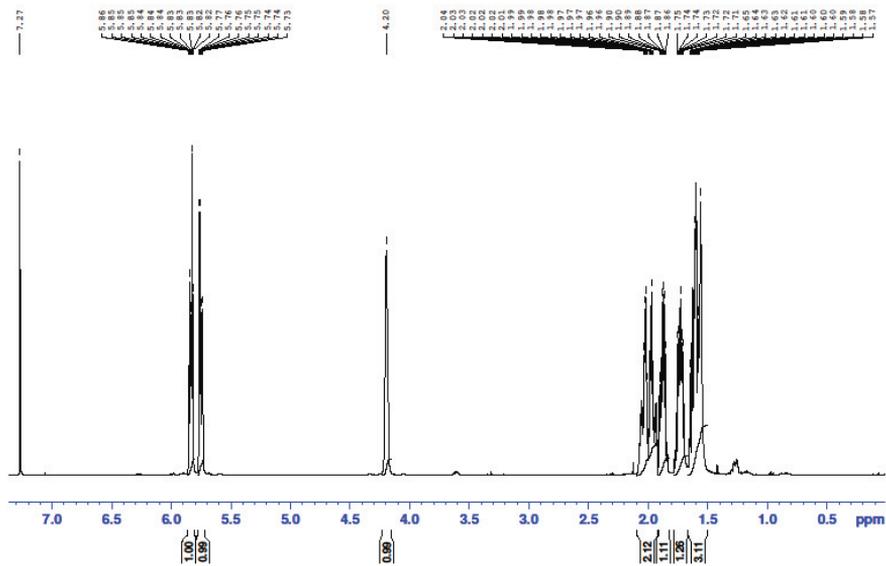
$^{13}\text{C}$  NMR 1-cyclohexylethane-1,2-diol (**2a**, AD-mix  $\beta$ )

**Figure S52:**  $^{13}\text{C}$  NMR of 1-cyclohexylethane-1,2-diol in  $\text{CDCl}_3$  isolated from the AD of vinylcyclohexane using AD-mix  $\beta$ .

HSQC 1-cyclohexylethane-1,2-diol (**2a**, AD-mix  $\beta$ )

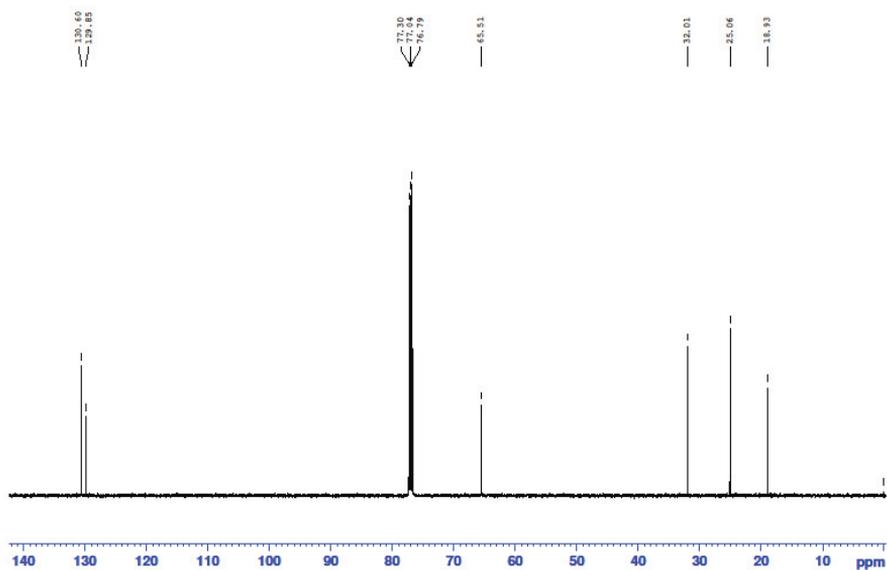
**Figure S53:** HSQC of 1-cyclohexylethane-1,2-diol in  $\text{CDCl}_3$  isolated from the AD of vinylcyclohexane using AD mix  $\beta$ .

$^1\text{H}$  NMR (*R*)-(+)-2-cyclohexenol ((*R*)-**3a**)



**Figure S54:**  $^1\text{H}$  NMR of (*R*)-(+)-2-cyclohexenol in  $\text{CDCl}_3$  isolated from the synthesis with  $\text{LiAlH}_4$  and cyclohexenone.

$^{13}\text{C}$  NMR (*R*)-(+)-2-cyclohexenol ((*R*)-**3a**)



**Figure S55:**  $^{13}\text{C}$  NMR of (*R*)-(+)-2-cyclohexenol in  $\text{CDCl}_3$  isolated from the synthesis with  $\text{LiAlH}_4$  and cyclohexenone.

$^1\text{H}$  NMR 1-indenol (**4a**, CDO)

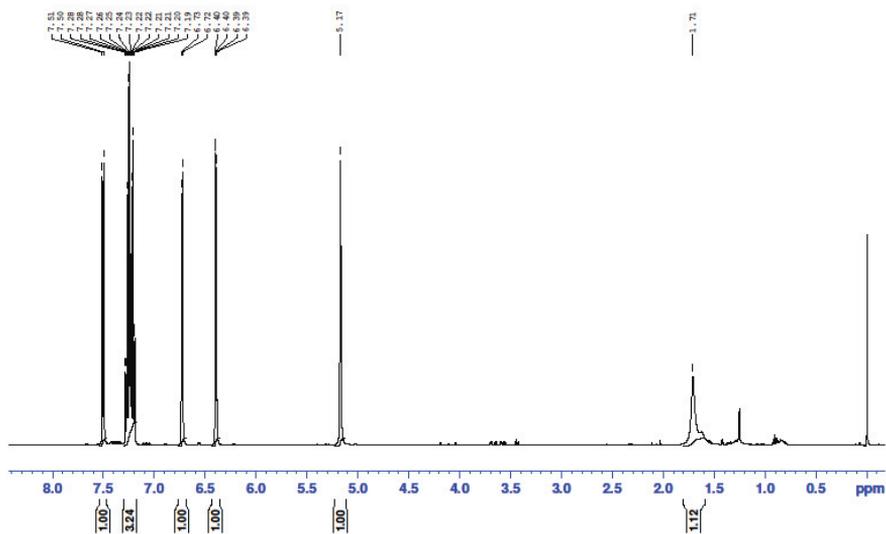
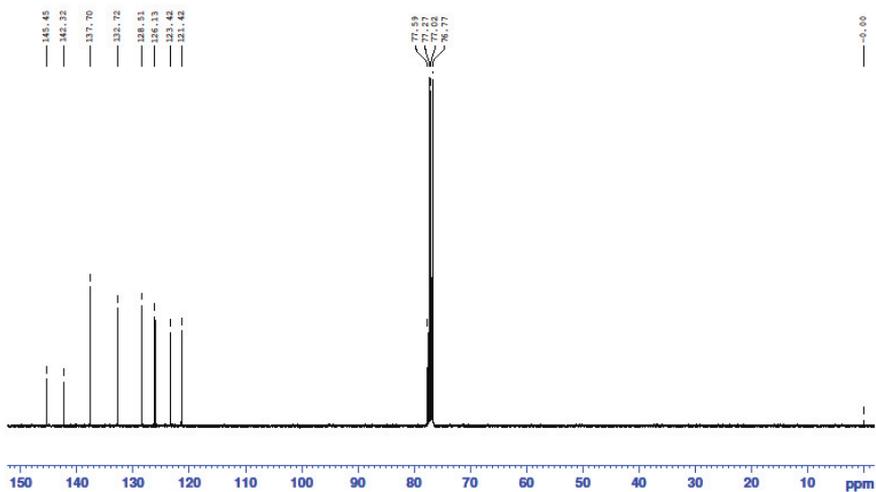


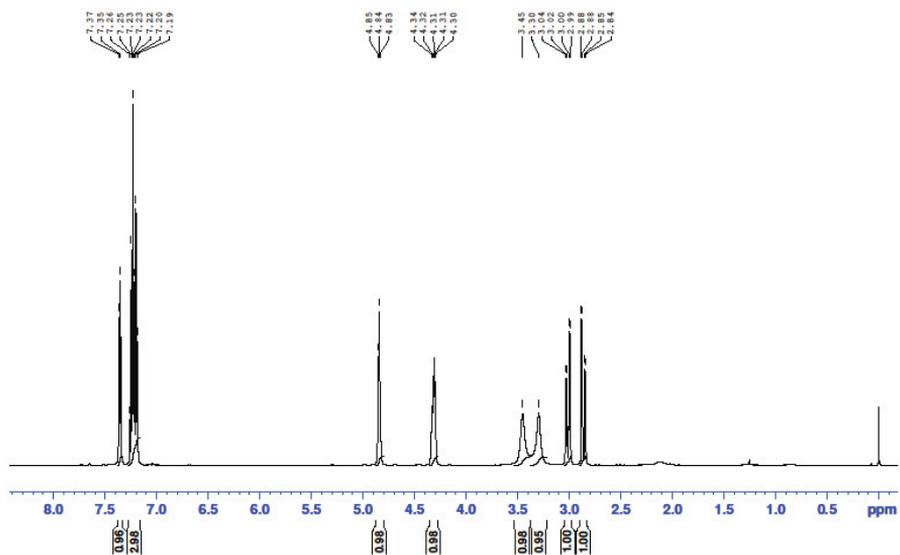
Figure S56:  $^1\text{H}$  NMR of 1-indenol in  $\text{CDCl}_3$  isolated from the CDO-catalyzed biotransformation of indene.

$^{13}\text{C}$  NMR 1-indenol (**4a**, CDO)



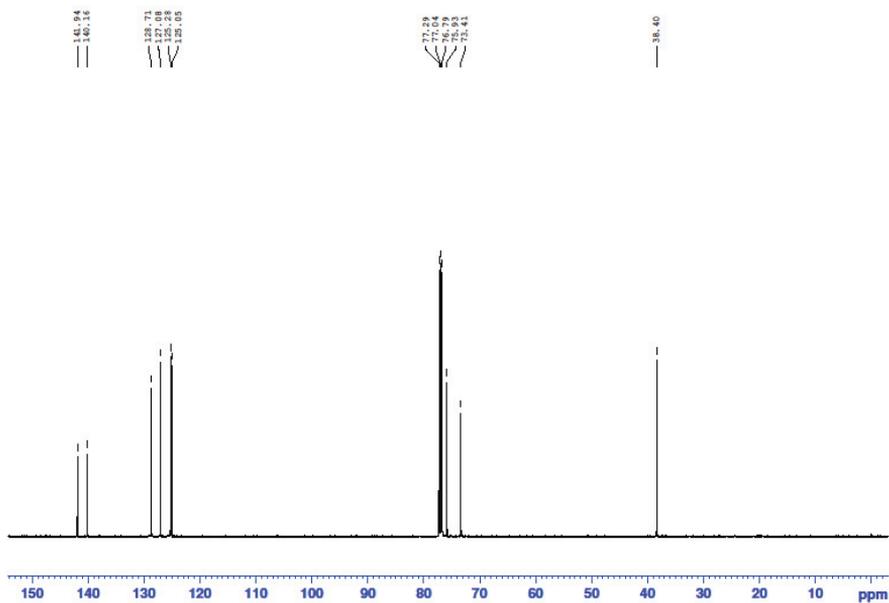
**Figure S57:**  $^{13}\text{C}$  NMR of 1-indenol in  $\text{CDCl}_3$  isolated from the CDO-catalyzed biotransformation of indene.

$^1\text{H}$  NMR *cis*-1,2-indandiol (**4b**, CDO)



**Figure S58:**  $^1\text{H}$  NMR of *cis*-1,2-indandiol in  $\text{CDCl}_3$  isolated from the CDO-catalyzed biotransformation of indene.

$^{13}\text{C}$  NMR *cis*-1,2-indandiol (**4b**, CDO)



**Figure S59:**  $^{13}\text{C}$  NMR of *cis*-1,2-indandiol in  $\text{CDCl}_3$  isolated from the CDO-catalyzed biotransformation of indene.

$^1\text{H}$  NMR *cis*-1,2-indandiol (**4b**, AD-mix  $\beta$ )

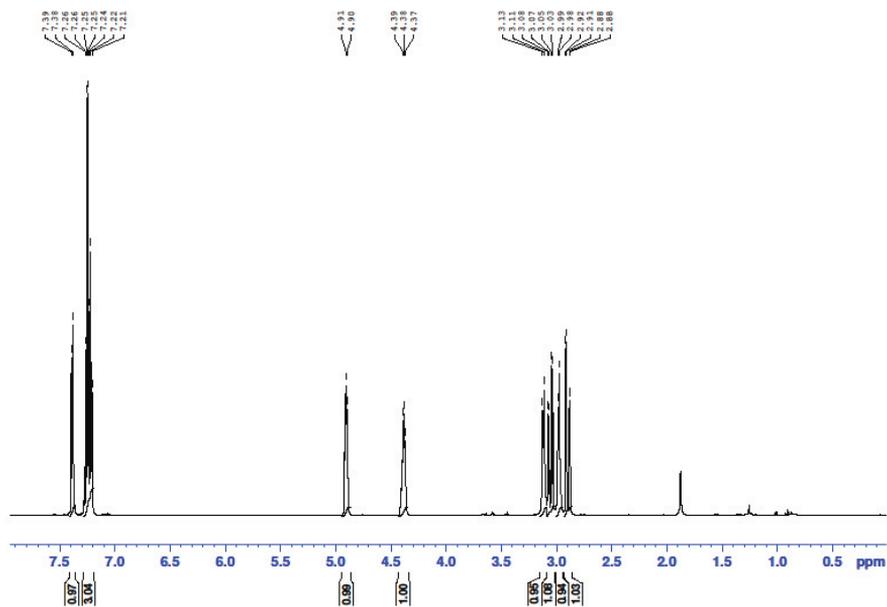
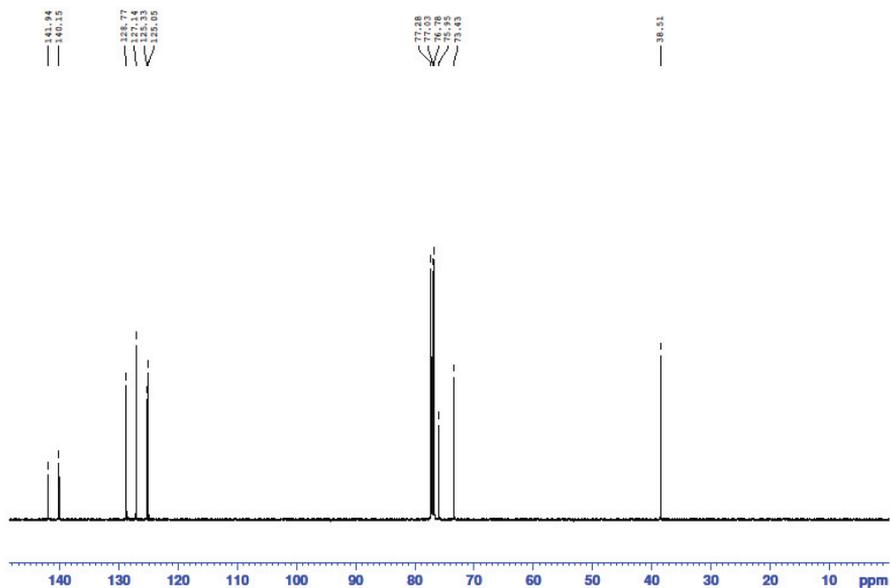


Figure S60:  $^1\text{H}$  NMR of *cis*-1,2-indandiol in  $\text{CDCl}_3$  isolated from the AD of indene using AD-mix  $\beta$ .

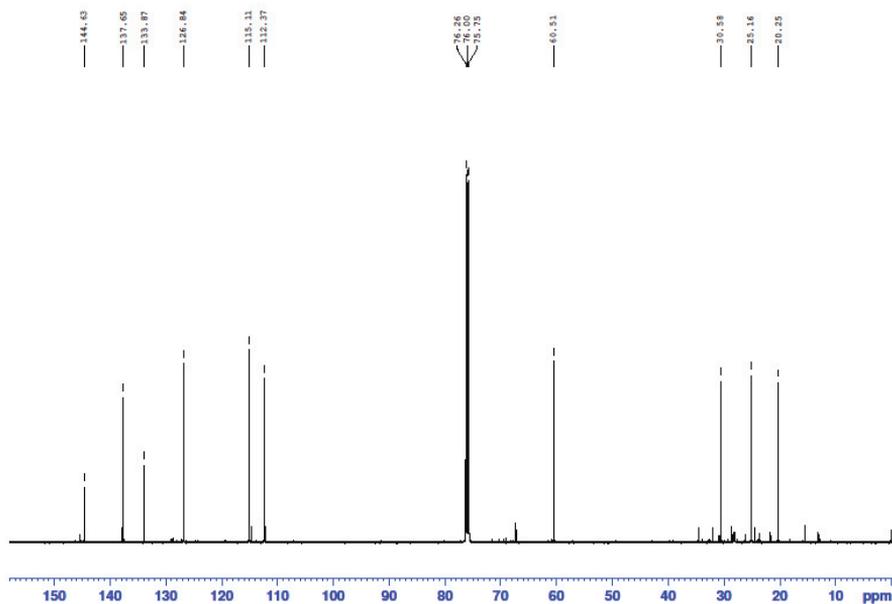
$^{13}\text{C}$  NMR *cis*-1,2-indandiol (**4b**, AD-mix  $\beta$ )



**Figure S61:**  $^{13}\text{C}$  NMR of *cis*-1,2-indandiol in  $\text{CDCl}_3$  isolated from the AD of indene using AD-mix  $\beta$ .

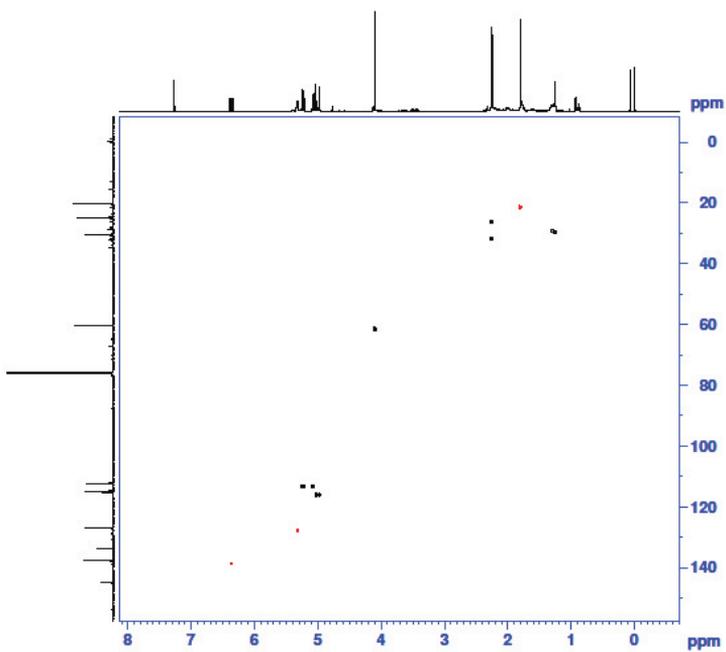


$^{13}\text{C}$  NMR (*Z*)-2-methyl-6-methyleneocta-2,7-dien-1-ol ((*Z*)-**5a**)



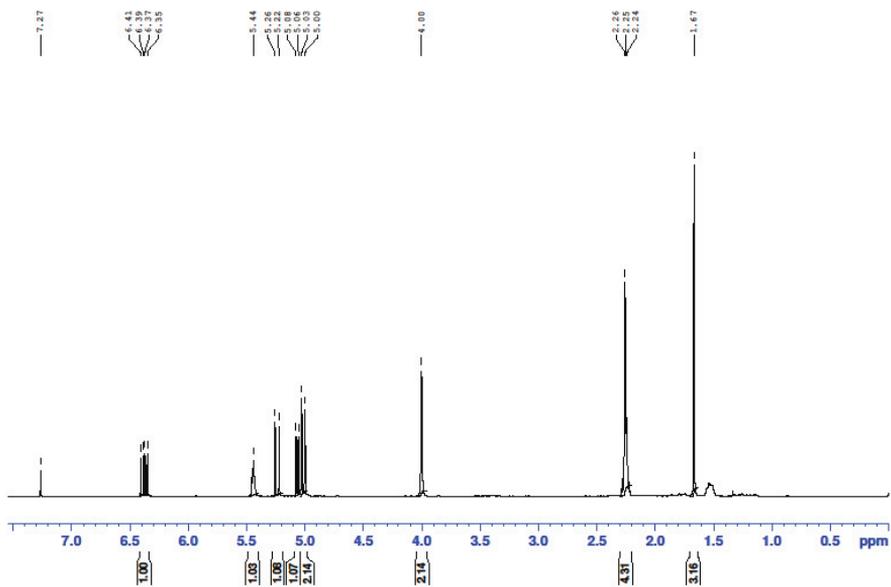
**Figure S63:**  $^{13}\text{C}$  NMR of (*Z*)-2-methyl-6-methyleneocta-2,7-dien-1-ol in  $\text{CDCl}_3$  isolated from the CDO M232A-catalyzed biotransformation of myrcene.

HSQC (*Z*)-2-methyl-6-methyleneocta-2,7-dien-1-ol ((*Z*)-**5a**)



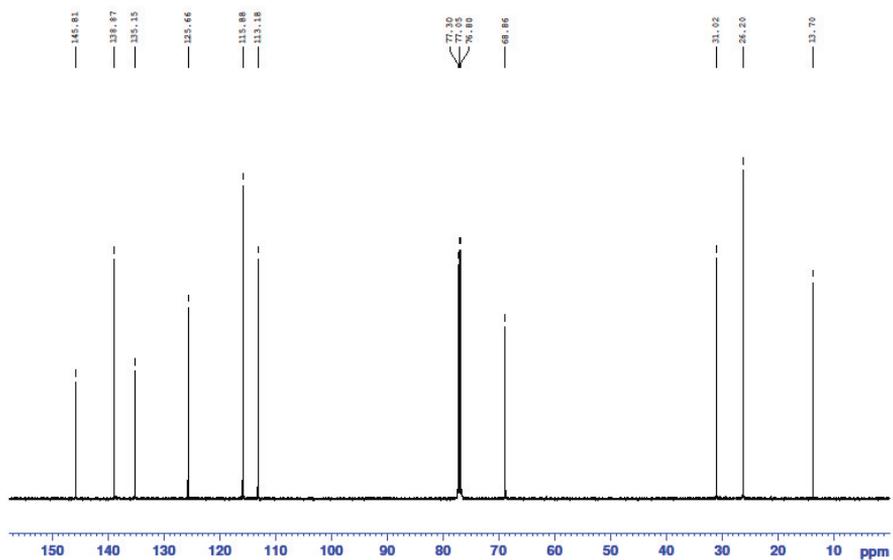
**Figure S64:** HSQC of (*Z*)-2-methyl-6-methyleneocta-2,7-dien-1-ol in CDCl<sub>3</sub> isolated from the CDO M232A-catalyzed biotransformation of myrcene.

$^1\text{H}$  NMR (*E*)-2-methyl-6-methyleneocta-2,7-dien-1-ol ((*E*)-**5a**)

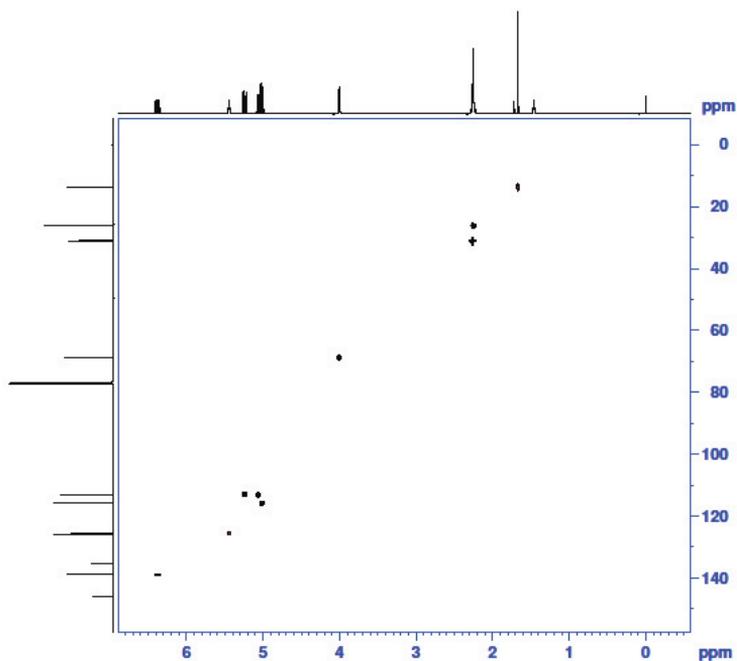


**Figure S65:**  $^1\text{H}$  NMR of (*E*)-2-methyl-6-methyleneocta-2,7-dien-1-ol in  $\text{CDCl}_3$  isolated from the synthesis with myrcene and  $\text{SeO}_2$ .

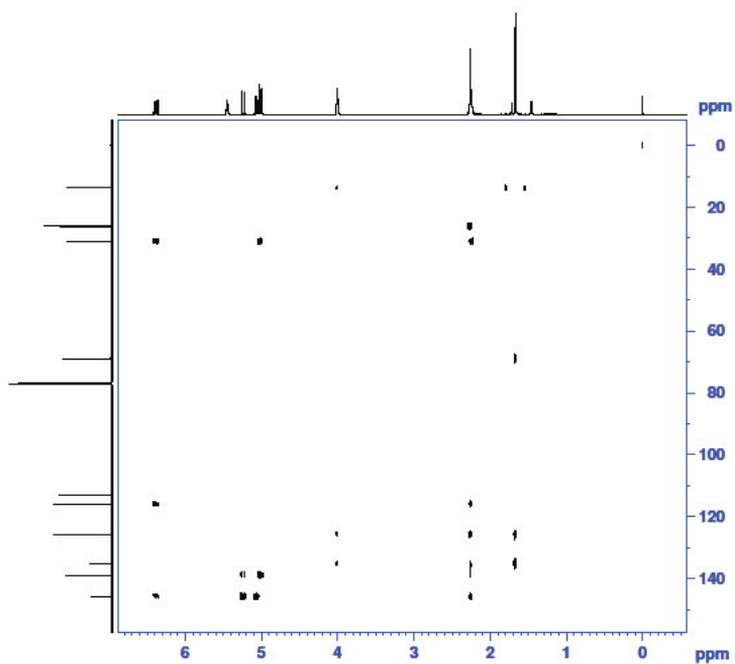
$^{13}\text{C}$  NMR (*E*)-2-methyl-6-methyleneocta-2,7-dien-1-ol ((*E*)-5a)



**Figure S66:**  $^{13}\text{C}$  NMR of (*E*)-2-methyl-6-methyleneocta-2,7-dien-1-ol in  $\text{CDCl}_3$  isolated from the synthesis with myrcene and  $\text{SeO}_2$ .

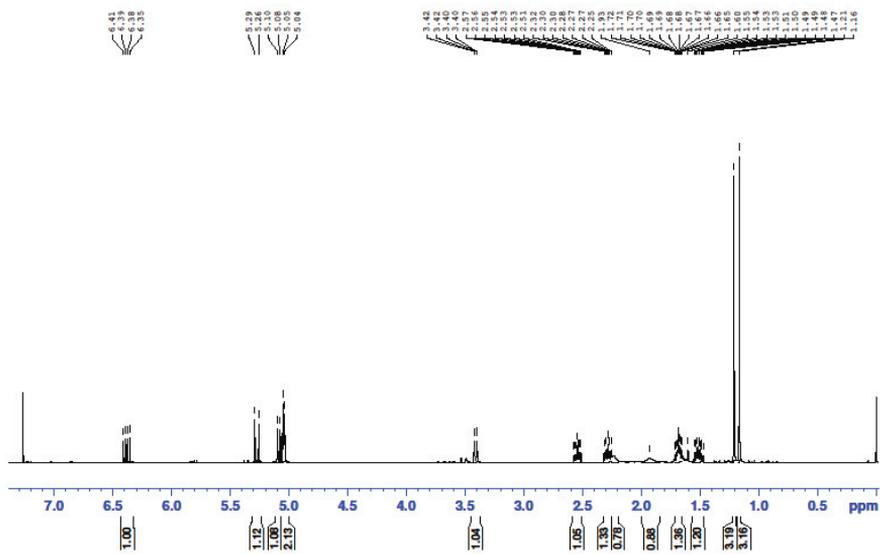
HSQC (*E*)-2-methyl-6-methyleneocta-2,7-dien-1-ol ((*E*)-5a)

**Figure S67:** HSQC of (*E*)-2-methyl-6-methyleneocta-2,7-dien-1-ol in  $\text{CDCl}_3$  isolated from the synthesis with myrcene and  $\text{SeO}_2$ .

HMBC (*E*)-2-methyl-6-methyleneocta-2,7-dien-1-ol ((*E*)-**5a**)

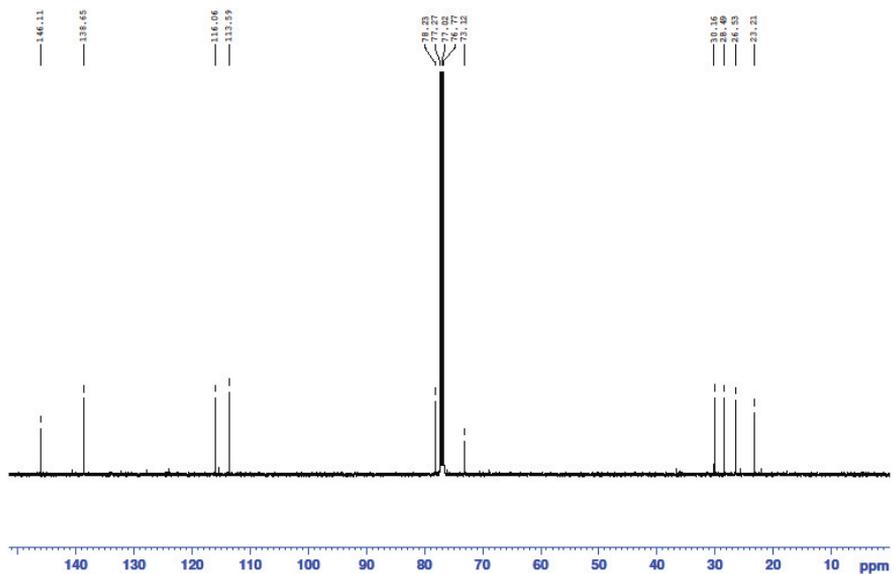
**Figure S68:** HMBC of (*E*)-2-methyl-6-methyleneocta-2,7-dien-1-ol in CDCl<sub>3</sub> isolated from the synthesis with myrcene and SeO<sub>2</sub>.

<sup>1</sup>H NMR 2-methyl-6-methyleneoct-7-ene-2,3-diol (**5c**, CDO M232A)



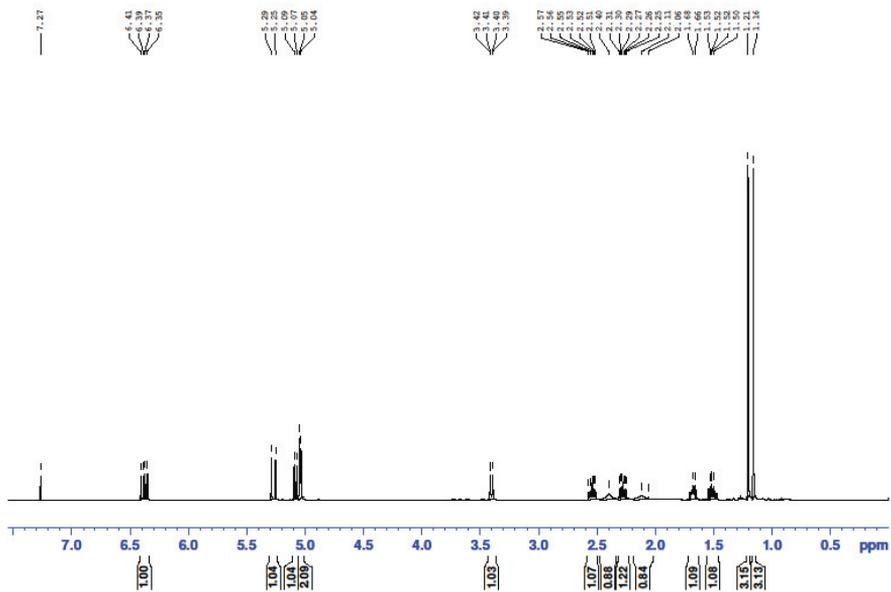
**Figure S69:** <sup>1</sup>H NMR of 2-methyl-6-methyleneoct-7-ene-2,3-diol in CDCl<sub>3</sub> isolated from the CDO M232A-catalyzed biotransformation of myrcene.

$^{13}\text{C}$  NMR 2-methyl-6-methyleneoct-7-ene-2,3-diol (**5c**, CDO M232A)



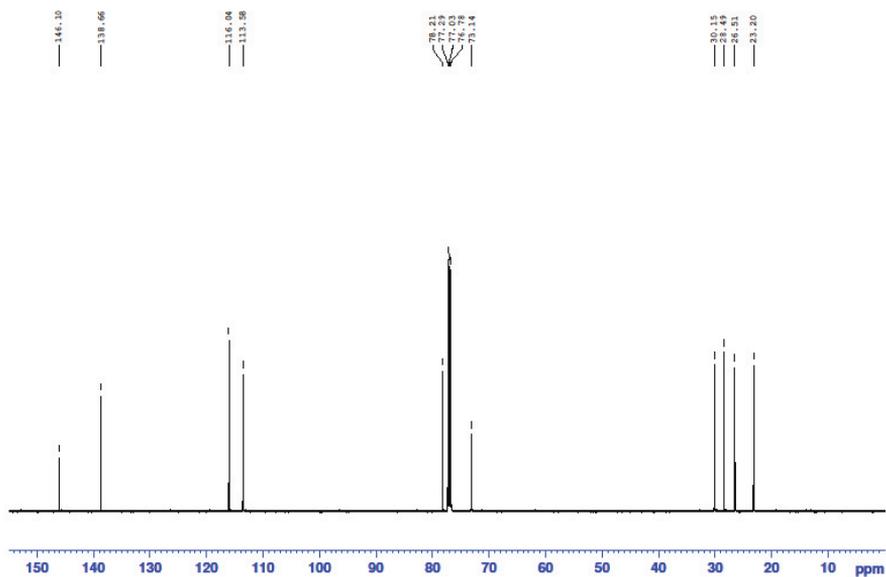
**Figure S70:**  $^{13}\text{C}$  NMR of 2-methyl-6-methyleneoct-7-ene-2,3-diol in  $\text{CDCl}_3$  isolated from the CDO M232A-catalyzed biotransformation of myrcene.

<sup>1</sup>H NMR 2-methyl-6-methyleneoct-7-ene-2,3-diol (**5c**, AD-mix β)



**Figure S71:** <sup>1</sup>H NMR of 2-methyl-6-methyleneoct-7-ene-2,3-diol in CDCl<sub>3</sub> isolated from the AD of myrcene using AD-mix β.

$^{13}\text{C}$  NMR 2-methyl-6-methyleneoct-7-ene-2,3-diol (**5c**, AD-mix  $\beta$ )



**Figure S72:**  $^{13}\text{C}$  NMR of 2-methyl-6-methyleneoct-7-ene-2,3-diol in  $\text{CDCl}_3$  isolated from the AD of myrcene using AD-mix  $\beta$ .

<sup>1</sup>H NMR 3,10-dihydro myrcene (**5b**)

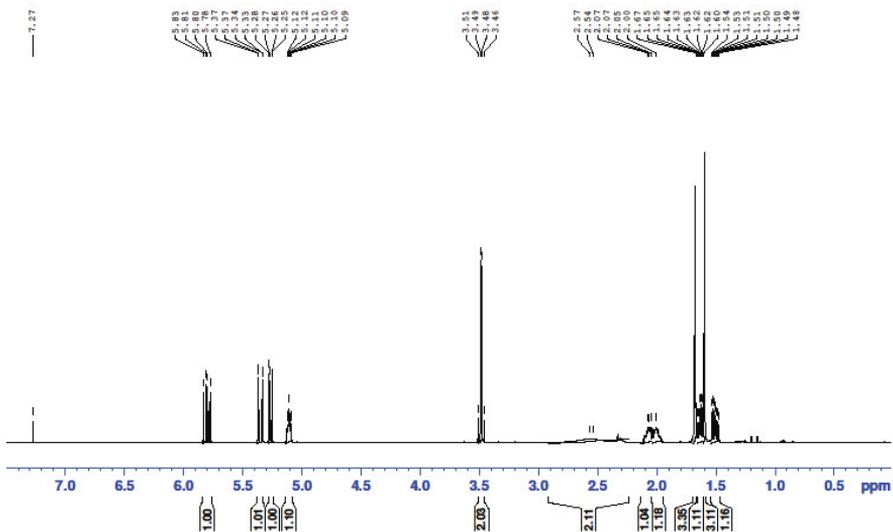
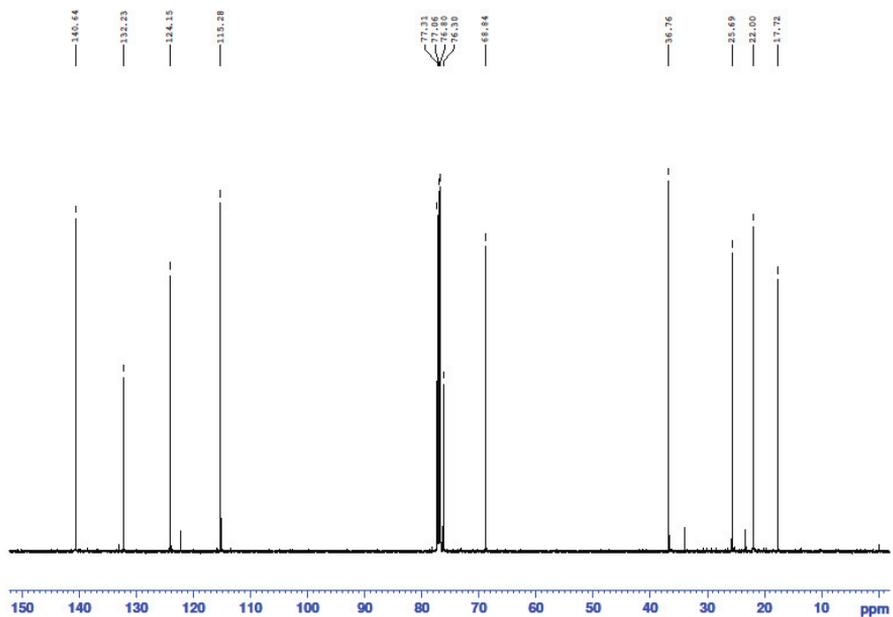
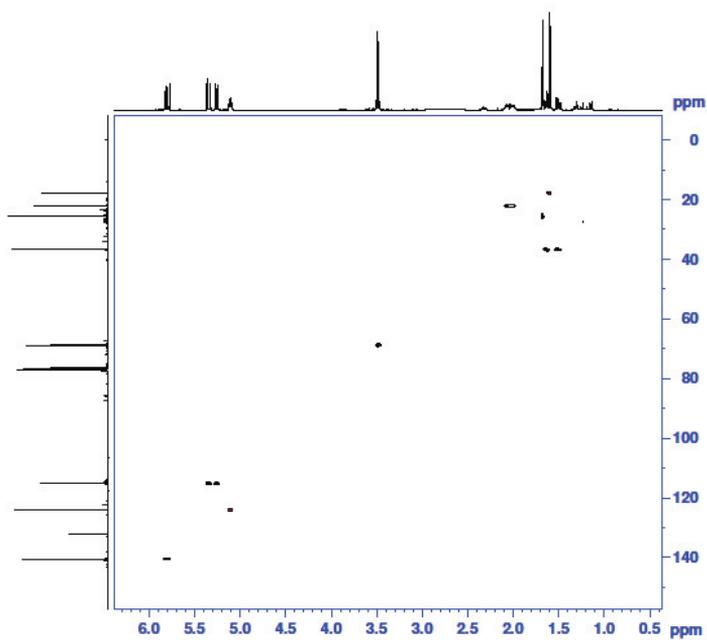


Figure S73: <sup>1</sup>H NMR of 3,10-dihydro myrcene in CDCl<sub>3</sub> isolated from the synthesis with myrcene and KMnO<sub>4</sub>.

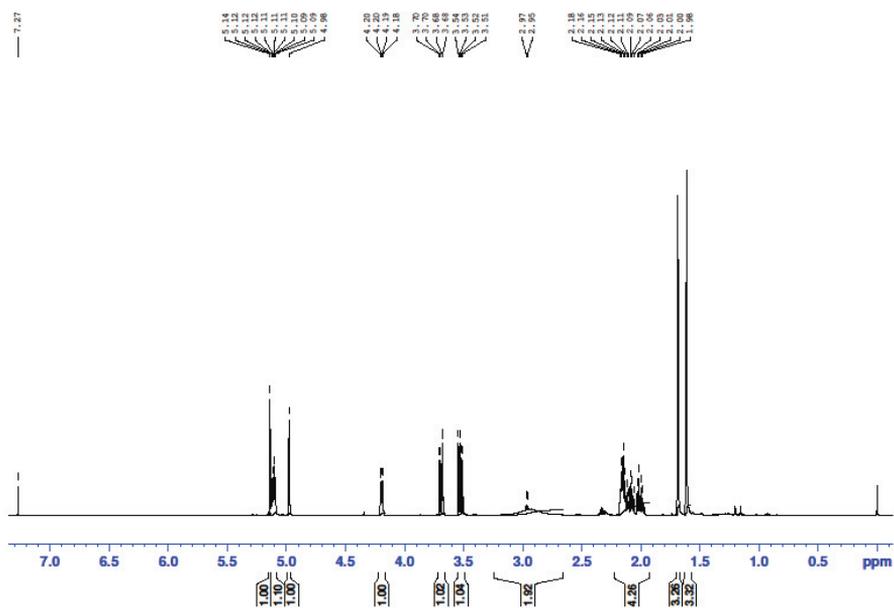
$^{13}\text{C}$  NMR 3,10-dihydro myrcene (**5b**)

**Figure S74:**  $^{13}\text{C}$  NMR of 3,10-dihydro myrcene in  $\text{CDCl}_3$  isolated from the synthesis with myrcene and  $\text{KMnO}_4$ .

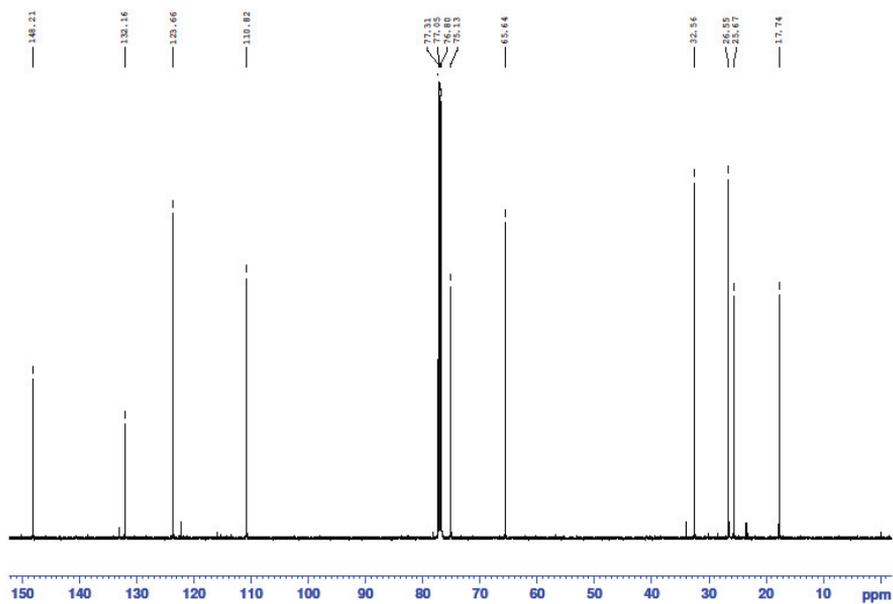
HSQC 3,10-dihydro myrcene (**5b**)

**Figure S75:** HSQC of 3,10-dihydro myrcene in  $\text{CDCl}_3$  isolated from the synthesis with myrcene and  $\text{KMnO}_4$ .

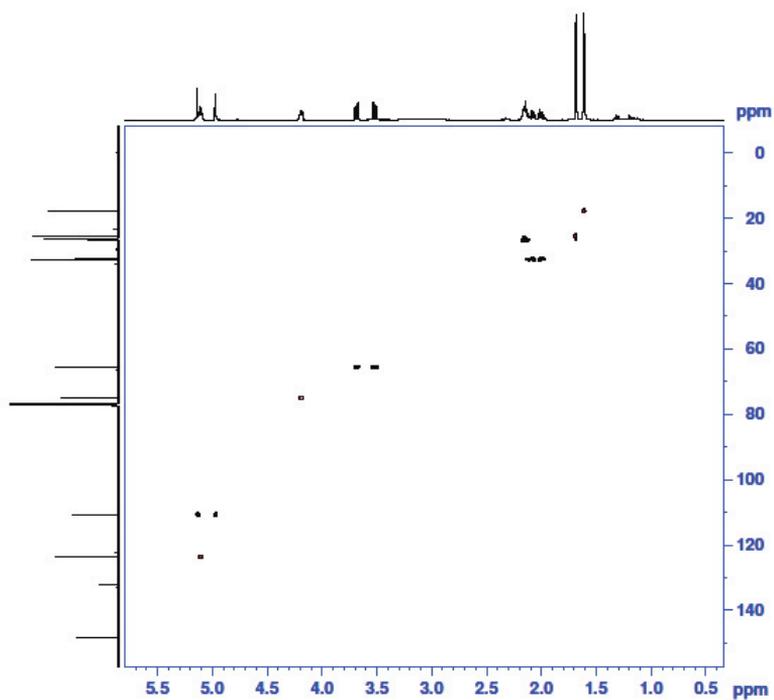
<sup>1</sup>H NMR 1,2-dihydro myrcene (**5d**)



**Figure S76:** <sup>1</sup>H NMR of 1,2-dihydro myrcene in CDCl<sub>3</sub> isolated from the synthesis with myrcene and KMnO<sub>4</sub>.

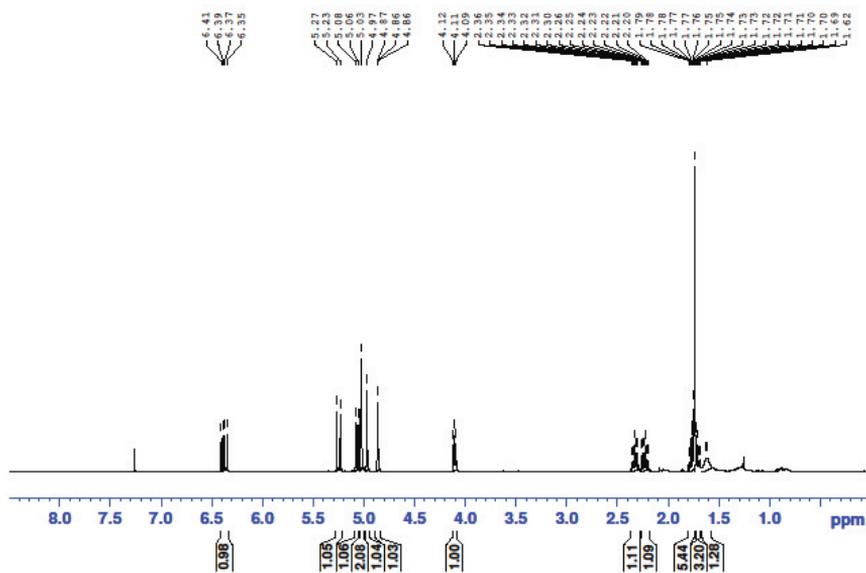
$^{13}\text{C}$  NMR 1,2-dihydro myrcene (**5d**)

**Figure S77:**  $^{13}\text{C}$  NMR of 1,2-dihydro myrcene in  $\text{CDCl}_3$  isolated from the synthesis with myrcene and  $\text{KMnO}_4$ .

HSQC 1,2-dihydro myrcene (**5d**)

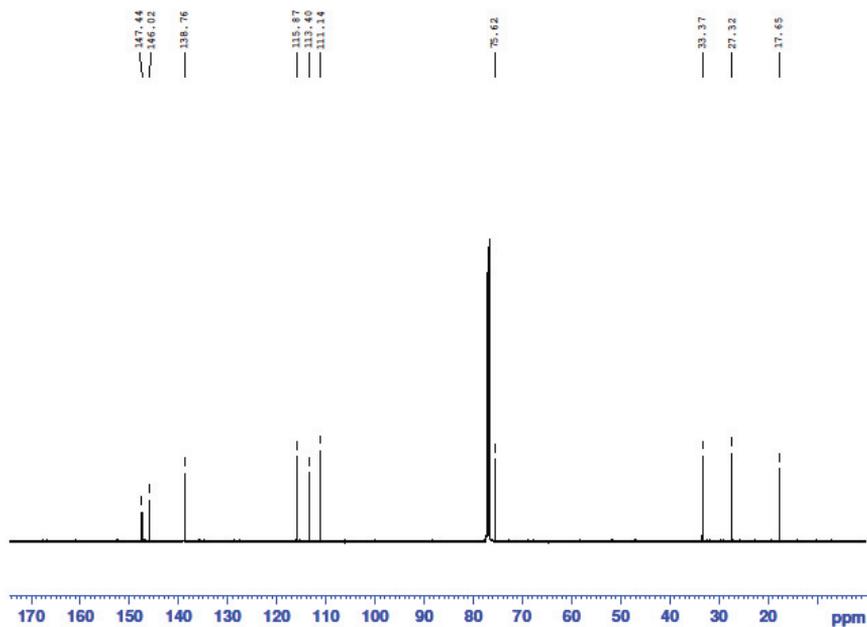
**Figure S78:** HSQC of 1,2-dihydro myrcene in  $\text{CDCl}_3$  isolated from the synthesis with myrcene and  $\text{KMnO}_4$ .

<sup>1</sup>H NMR 2-methyl-6-methylene-octa-1,7-dien-3-ol (CDO M232A)



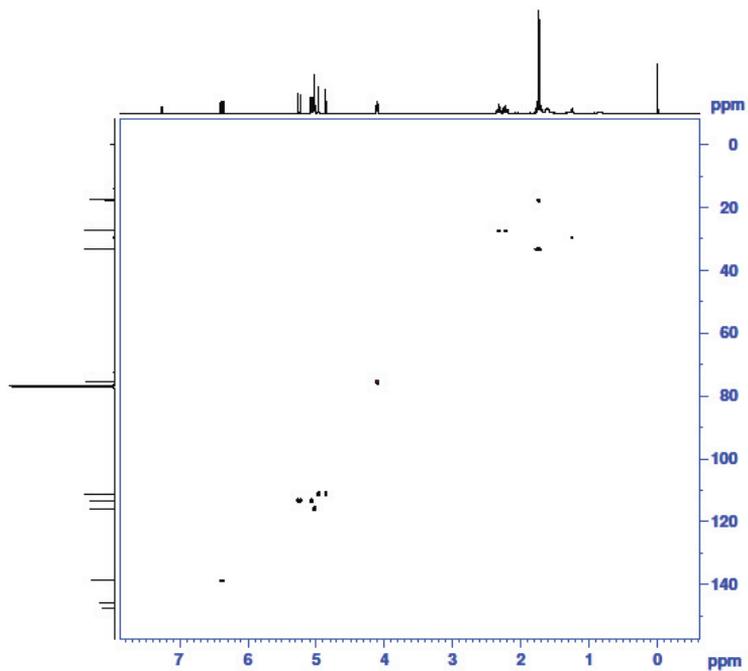
**Figure S79:** <sup>1</sup>H NMR of 2-methyl-6-methylene-octa-1,7-dien-3-ol in CDCl<sub>3</sub> isolated from the CDO M232A-catalyzed biotransformation of myrcene.

$^{13}\text{C}$  NMR 2-methyl-6-methylene-octa-1,7-dien-3-ol (CDO M232A)

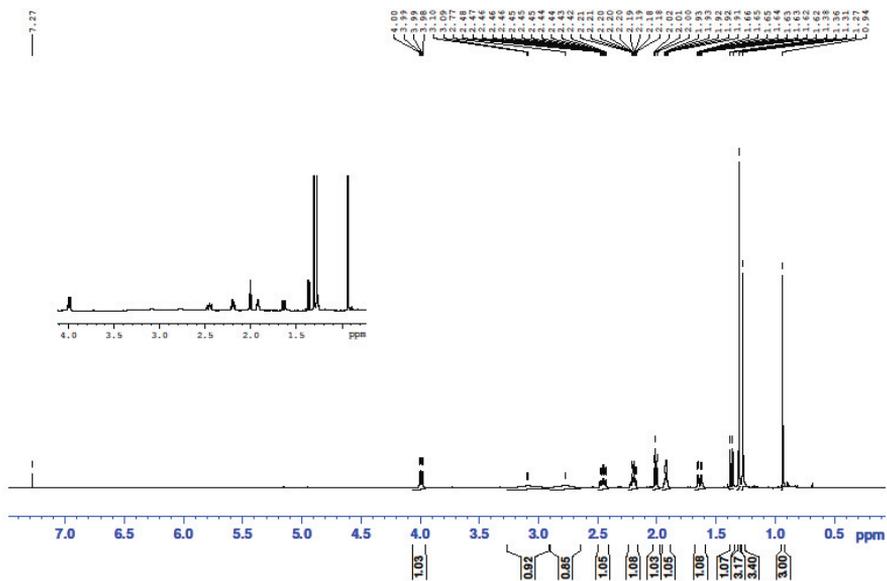


**Figure S80:**  $^{13}\text{C}$  NMR of 2-methyl-6-methylene-octa-1,7-dien-3-ol in  $\text{CDCl}_3$  isolated from the CDO M232A-catalyzed biotransformation of myrcene.

## HSQC 2-methyl-6-methylene-octa-1,7-dien-3-ol (CDO M232A)

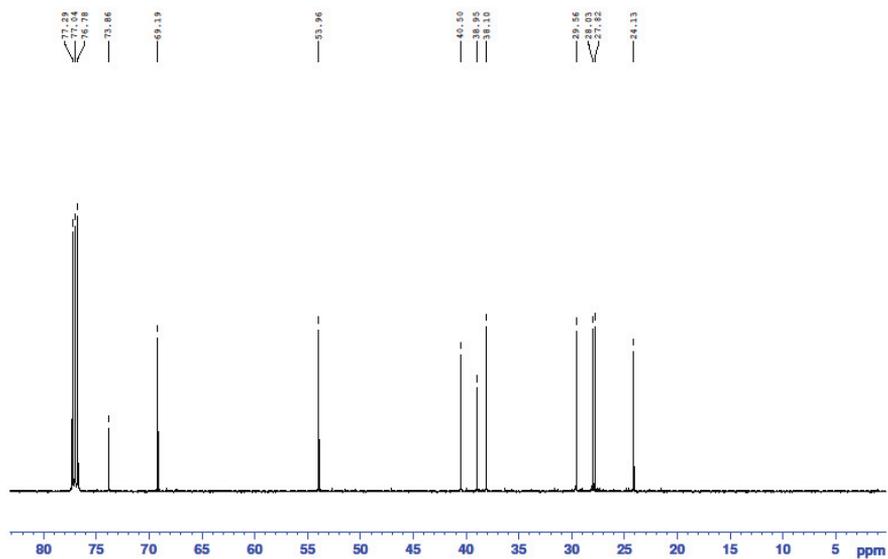


**Figure S81:** HSQC of 2-methyl-6-methylene-octa-1,7-dien-3-ol in  $\text{CDCl}_3$  isolated from the CDO M232A-catalyzed biotransformation of myrcene.

$^1\text{H}$  NMR pinanediol (**6a**, CDO M232A)


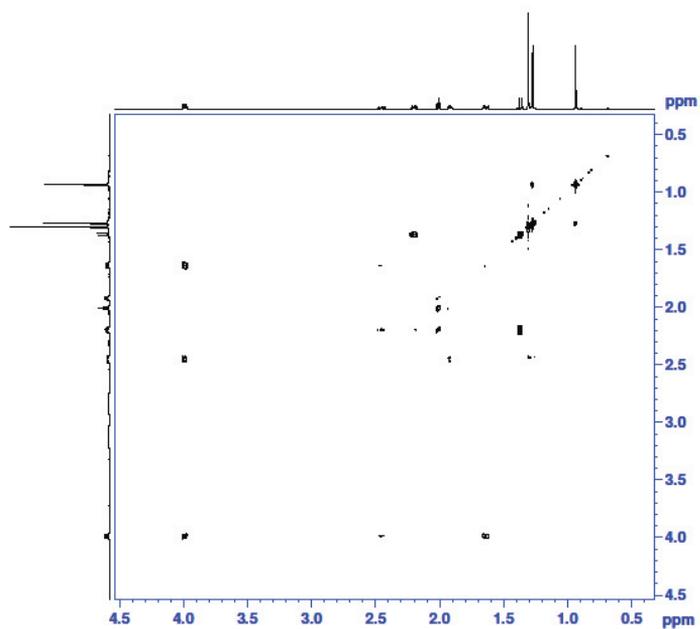
**Figure S82:**  $^1\text{H}$  NMR of pinanediol in  $\text{CDCl}_3$  isolated from the CDO M232A-catalyzed biotransformation of (+)- $\alpha$ -pinene.

$^{13}\text{C}$  NMR pinanediol (**6a**, CDO M232A)



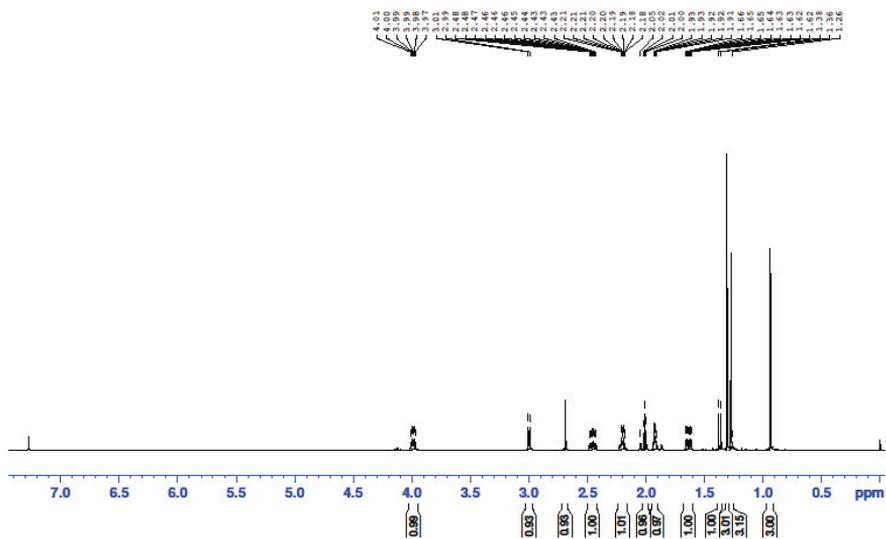
**Figure S83:**  $^{13}\text{C}$  NMR of pinanediol in  $\text{CDCl}_3$  isolated from the CDO M232A-catalyzed biotransformation of (+)- $\alpha$ -pinene.

$^1\text{H}$   $^1\text{H}$  COSY pinanediol (**6a**, CDO M232A)



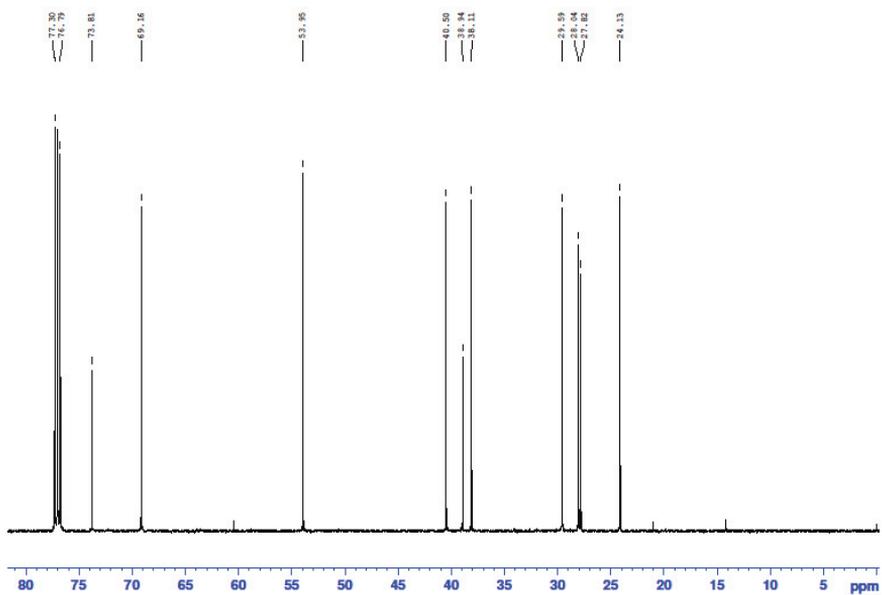
**Figure S84:**  $^1\text{H}$   $^1\text{H}$  COSY of pinanediol in  $\text{CDCl}_3$  isolated from the CDO M232A-catalyzed biotransformation of (+)- $\alpha$ -pinene.

<sup>1</sup>H NMR pinanediol (**6a**, OsO<sub>4</sub>)

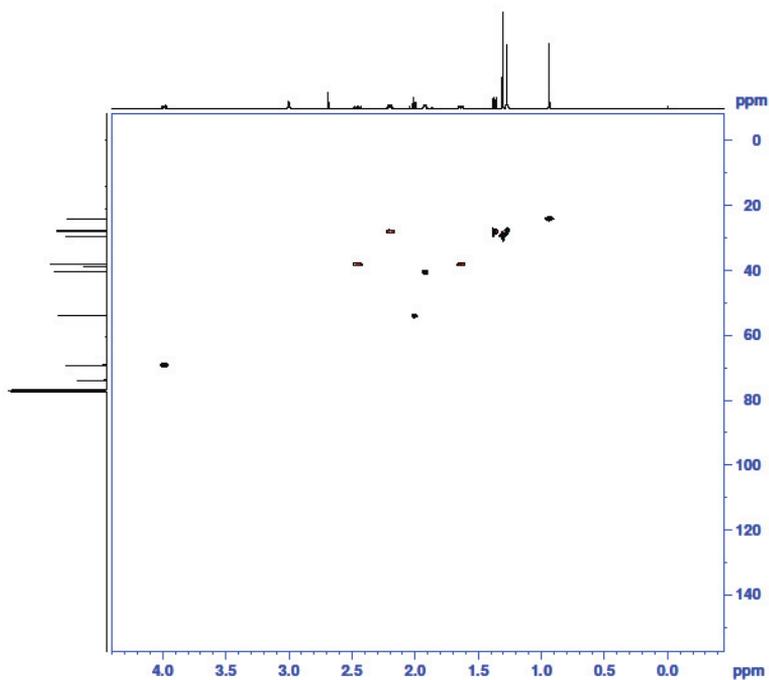


**Figure S85:** <sup>1</sup>H NMR of pinanediol in CDCl<sub>3</sub> isolated from the synthesis with OsO<sub>4</sub> and (+)- $\alpha$ -pinene.

$^{13}\text{C}$  NMR pinanediol (**6a**,  $\text{OsO}_4$ )

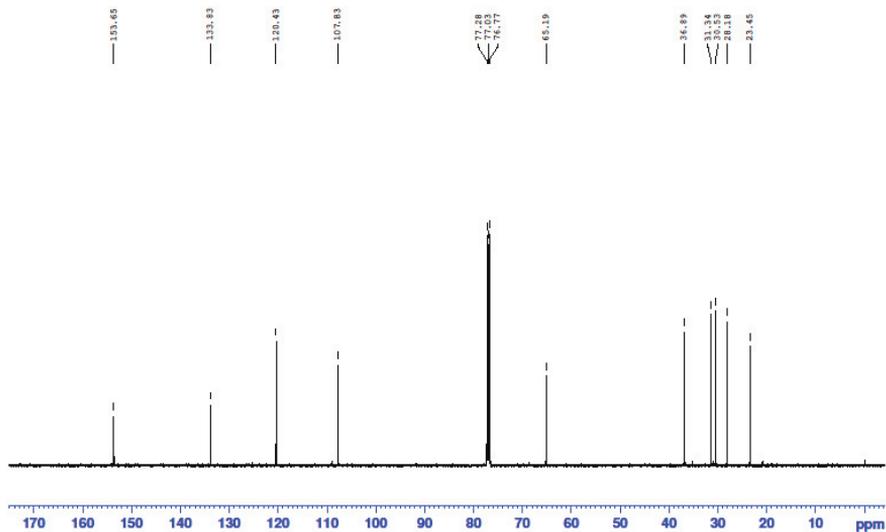


**Figure S86:**  $^{13}\text{C}$  NMR of pinanediol in  $\text{CDCl}_3$  isolated from the synthesis with  $\text{OsO}_4$  and (+)- $\alpha$ -pinene.

HSQC pinanediol (**6a**, OsO<sub>4</sub>)

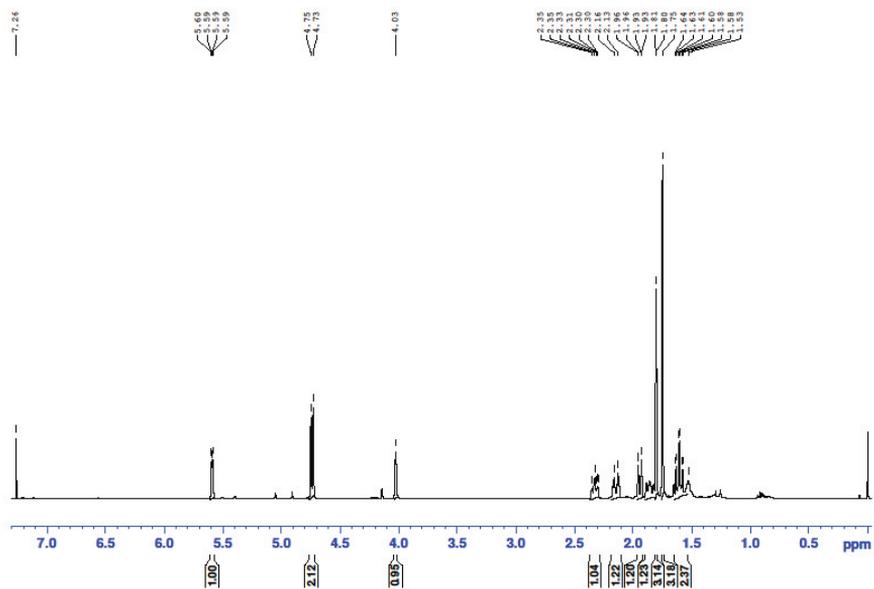
**Figure S87:** HSQC of pinanediol in CDCl<sub>3</sub> isolated from the synthesis with OsO<sub>4</sub> and (+)- $\alpha$ -pinene.



$^{13}\text{C}$  NMR (+)-mentha-1.8-dien-10-ol (**7a**)

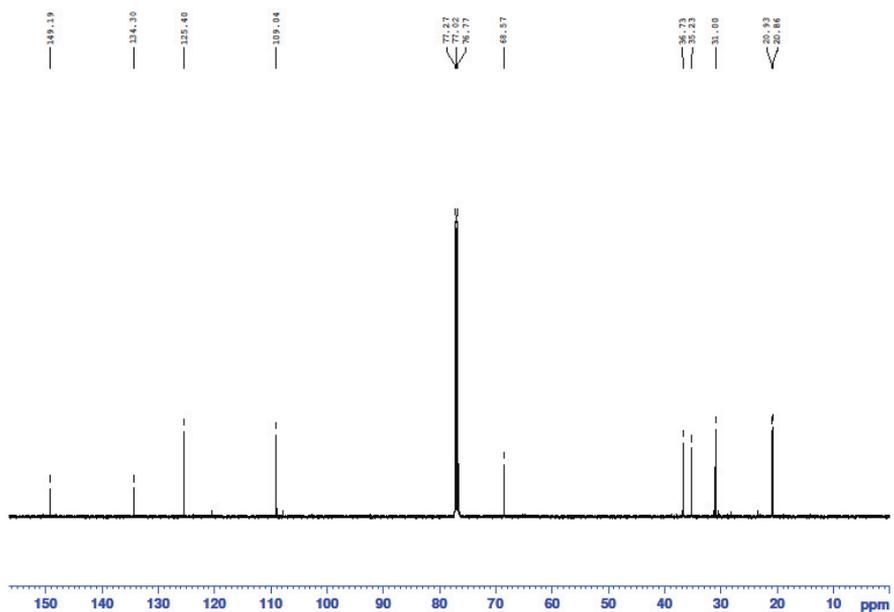
**Figure S89:**  $^{13}\text{C}$  NMR of (+)-mentha-1.8-dien-10-ol in  $\text{CDCl}_3$  isolated from the synthesis with  $\text{SeO}_2$  and (*R*)-limonene.

$^1\text{H}$  NMR carveol (**7b**, CDO M232A)



**Figure S90:**  $^1\text{H}$  NMR of carveol in  $\text{CDCl}_3$  isolated from the CDO M232A-catalyzed biotransformation of (*R*)-limonene.

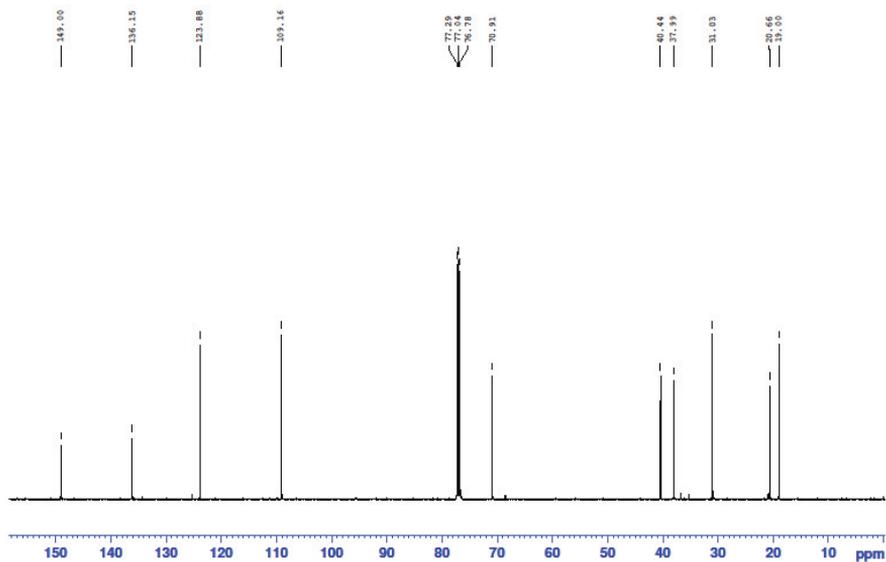
<sup>13</sup>C NMR carveol (7b, CDO M232A)



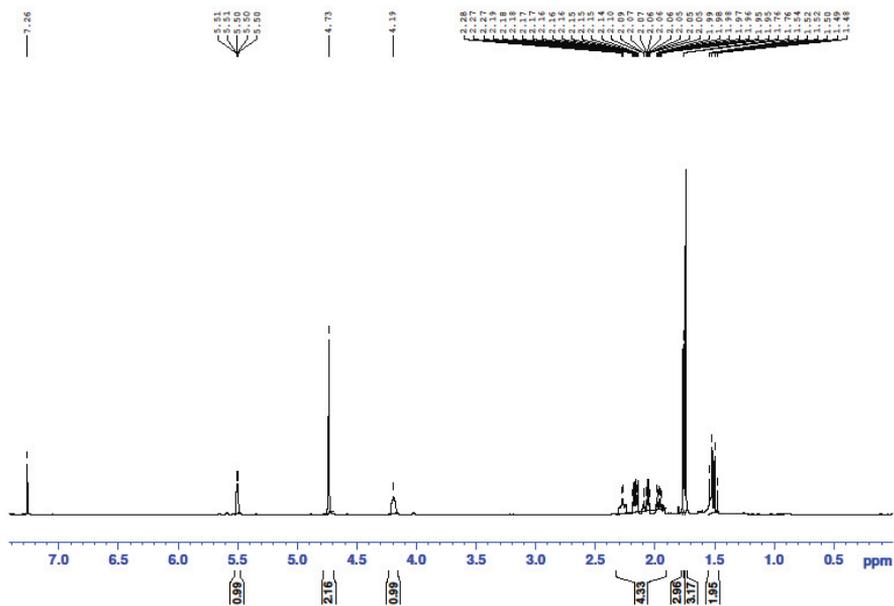
**Figure S91:** <sup>13</sup>C NMR of carveol in CDCl<sub>3</sub> isolated from the CDO M232A-catalyzed biotransformation of (*R*)-limonene.



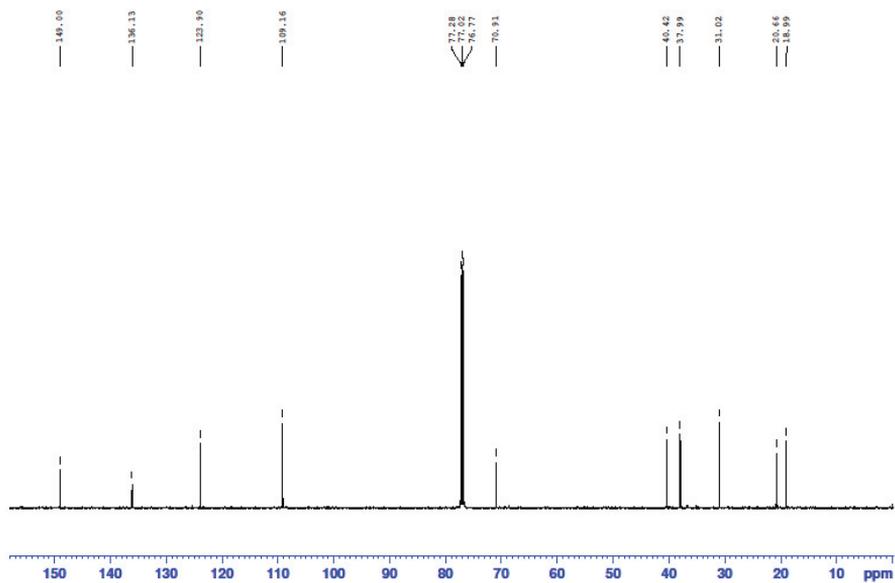
$^{13}\text{C}$  NMR (1*S*,5*S*)-carveol ((1*S*,5*S*)-**7b**,  $\text{LiAlH}_4$ )



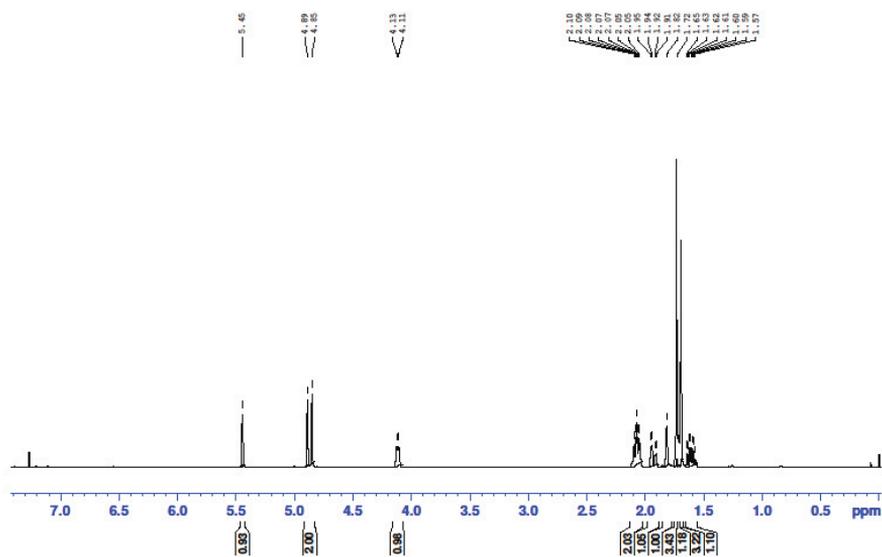
**Figure S93:**  $^{13}\text{C}$  NMR of (1*S*,5*S*)-carveol in  $\text{CDCl}_3$  isolated from the synthesis with  $\text{LiAlH}_4$  and (*S*)-(+)-carvone.

$^1\text{H}$  NMR (*1R,5R*)-carveol ((*1R,5R*)-**7b**)

**Figure S94:**  $^1\text{H}$  NMR of (*1R,5R*)-carveol in  $\text{CDCl}_3$  isolated from the synthesis with  $\text{LiAlH}_4$  and (*R*)-(-)-carvone.

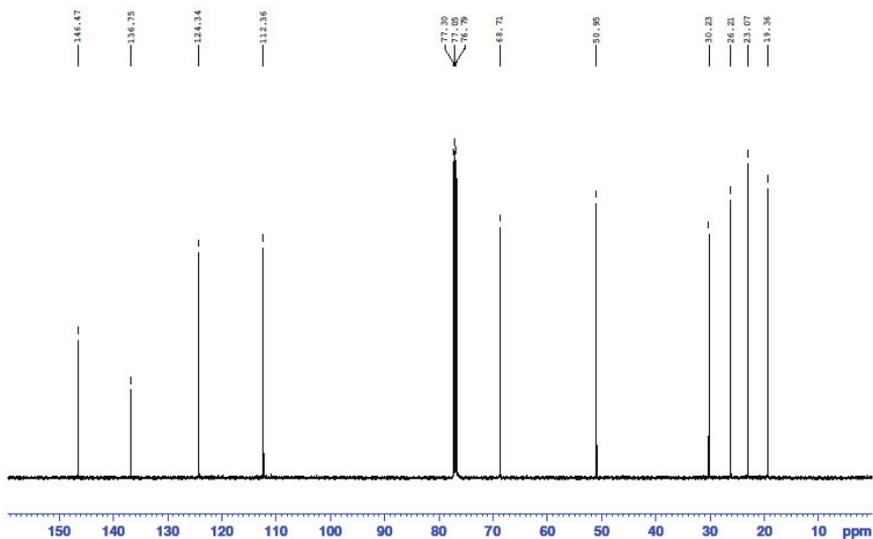
$^{13}\text{C}$  NMR (1*R*,5*R*)-carveol ((1*R*,5*R*)-**7b**)

**Figure S95:**  $^{13}\text{C}$  NMR of (1*R*,5*R*)-carveol in  $\text{CDCl}_3$  isolated from the synthesis with  $\text{LiAlH}_4$  and (*R*)-(-)-carvone.

$^1\text{H}$  NMR isopiperitenol (**7c**)

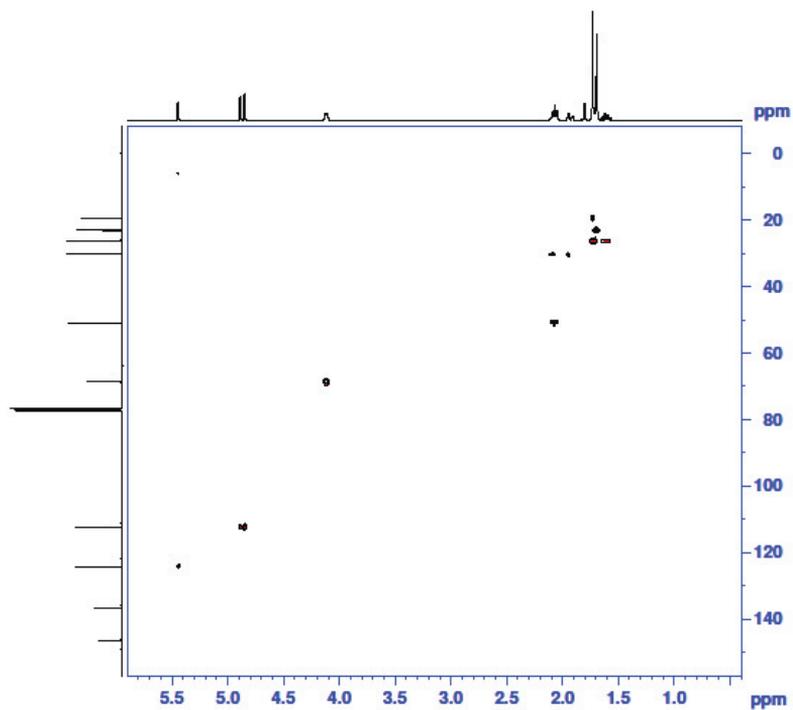
**Figure S96:**  $^1\text{H}$  NMR of isopiperitenol in  $\text{CDCl}_3$  isolated from the CDO I288V-catalyzed biotransformation of (*S*)-limonene.

$^{13}\text{C}$  NMR isopiperitenol (**7c**)



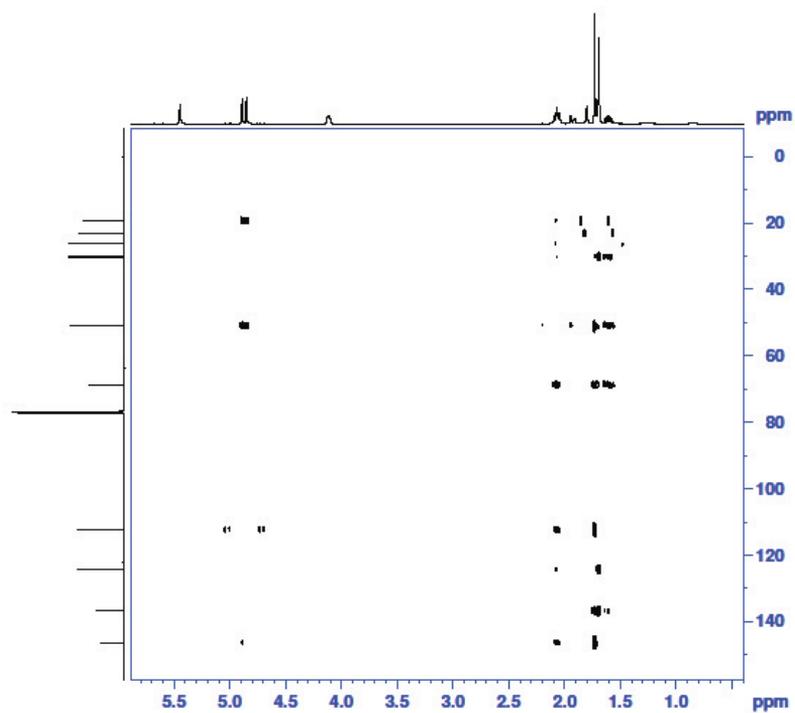
**Figure S97:**  $^{13}\text{C}$  NMR of isopiperitenol in  $\text{CDCl}_3$  isolated from the CDO I288V-catalyzed biotransformation of (*S*)-limonene.

## HSQC isopiperitenol (7c)

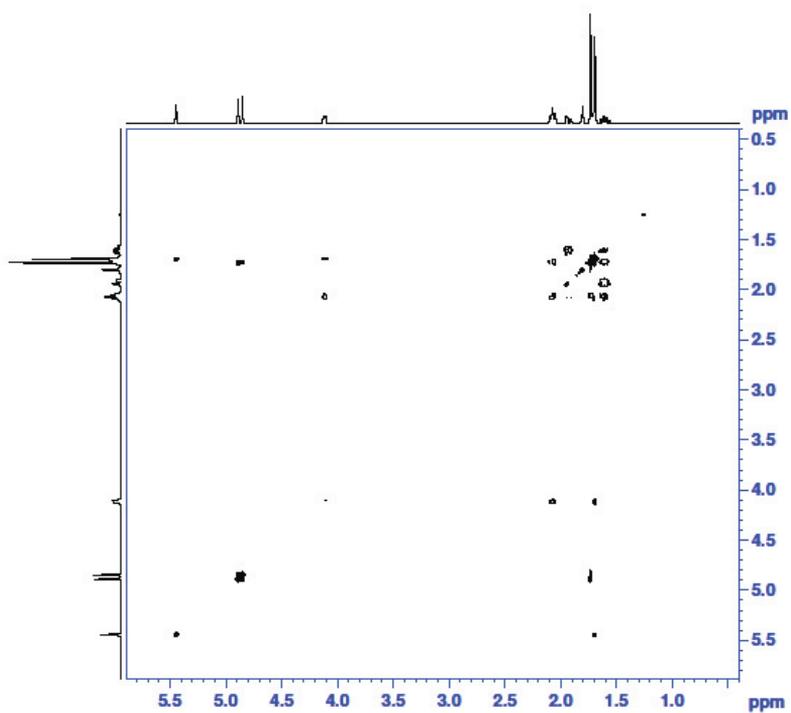


**Figure S98:** HSQC of isopiperitenol in CDCl<sub>3</sub> isolated from the CDO I288V-catalyzed biotransformation of (*S*)-limonene.

## HMBC isopiperitenol (7c)



**Figure S99:** HMBC of isopiperitenol in CDCl<sub>3</sub> isolated from the CDO I288V-catalyzed biotransformation of (*S*)-limonene.

COSY isopiperitenol (**7c**)

**Figure S100:** COSY of isopiperitenol in  $\text{CDCl}_3$  isolated from the CDO I288V-catalyzed biotransformation of (*S*)-limonene.

## 8. CURRICULUM VITAE

### PERSÖNLICHE DATEN

---

Name Christine Gally  
Geburtsdaten 10.07.1985 in Karlsruhe  
Nationalität Deutsch

### AKADEMISCHE UND SCHULISCHE LAUFBAHN

---

- ab 01/2012**      **Promotion, Institut für Technische Biochemie, Abteilung Biokatalyse, Universität Stuttgart**  
Thema: Enzymatische asymmetrische Dihydroxylierung von Alkenen,  
Betreuer: Prof. Dr. Bernhard Hauer
- 10/2005-11/2011**      **Studium Technische Biologie, Universität Stuttgart**  
Abschluss: Diplom-Biologin t.o. (technisch orientiert)  
Note: 1,0 (mit Auszeichnung bestanden)
- 10/2010-07/2011**      Diplomarbeit: „*Effect of detergents on substrate conversion by squalene hopene cyclases*“, Institut für Technische Biochemie, Universität Stuttgart,  
Betreuer: Prof. Dr. Bernhard Hauer
- 10/2009-06/2010**      Studienarbeit: „*Activation-induced cytidine deaminase in autoimmunity*“, Department of Microbiology & Immunology, University of California, San Francisco,  
Betreuer: Prof. Dr. Matthias Wabl
- 09/1996-06/2005**      **Albertus-Magnus-Gymnasium, Ettlingen**  
Abschluss: Allgemeine Hochschulreife

### PRAKTISCHE ERFAHRUNG

---

- 09/2015-11/2015**      **Orion Pharma, Abteilung Medizinische Chemie, Espoo Finland**  
Technologietransfer im Bereich Biokatalyse
- 07/2011-10/2011**      **Tosoh Bioscience GmbH, Stuttgart**  
Praktikum im Bereich Proteinaufreinigung; Methodenoptimierung für die biochromatographische Aufreinigung mono- und polyklonaler Antikörper
- 03/2009-10/2009**      **Wissenschaftliche Hilfskraft, Universität Stuttgart**  
Proteinengineering und Charakterisierung von P450 BM3 Mutanten