Studies of the toluene dioxygenase from *Pseudomonas putida* F1: influence of active-site positions on hydroxylations of mono- and bicyclic aromatics

Studien zur Toluoldioxygenase aus *Pseudomonas putida* F1: Einfluss von Positionen des aktiven Zentrums auf Hydroxylierungen von mono- und bizyklischen Aromaten

Von der Fakultät 3 (Chemie) der Universität Stuttgart

zur Erlangung der Würde eines Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigte Abhandlung

Vorgelegt von

Julian Lewin Wissner

aus Stuttgart

Hauptberichter: Prof. Dr. Bernhard Hauer

Mitberichter: Prof. Dr. René Peters

Vorsitzender: Prof. Dr. Albert Jeltsch

Tag der mündlichen Prüfung: 19.10.2021

Institut für Biochemie und Technische Biochemie der Universität Stuttgart

Abteilung Technische Biochemie

- 2021 -

DECLARATION OF AUTHORSHIP

I hereby certify that the dissertation entitled "Studies of the toluene dioxygenase from *Pseudomonas putida* F1: influence of active-site positions on hydroxylations of mono- and bicyclic aromatics" is entirely my own work except where otherwise indicated. Passages and ideas from other sources have been clearly indicated. Furthermore, I confirm that the thesis presented here has not been submitted in the same or similar form to any other institution for the purpose of obtaining an academic degree.

ERKLÄRUNG ÜBER DIE EIGENSTÄNDIGKEIT DER DISSERTATION

Ich versichere, dass ich die vorliegende Arbeit mit dem Titel „Studien zur Toluoldioxygenase aus *Pseudomonas putida* F1: Einfluss von Positionen des aktiven Zentrums auf Hydroxylierungen von mono- und bizyklischen Aromaten" 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, dass die hier vorgelegte Thesis nicht in gleicher oder ähnlicher Form bei einer andere Institution zur Erlangung eines akademischen Grades eingereicht wurde.

Name/Name: Julian Lewin Wissner

Signed/Unterschrift: _____

Date/Datum: 14.06.2021

The present doctoral thesis was prepared under the direction and supervision of Prof. Dr. Bernhard Hauer at the Institute of Biochemistry and Technical Biochemistry, Department of Technical Biochemistry at the University of Stuttgart, Germany. The experimental research was conducted in the period from January 2018 to March 2021 and was funded by the BMBF as part of the project *PowerCart* (Grant Agreement No. 031B0369A). The content of this thesis, in cumulative format, is a comprehensive compendium of three published scientific articles generated during the course of the doctoral work, which are identified with the numbers 2, 4, and 5 in the publication list, and correspond to articles I, II and III.

PUBLICATION LIST (2020-2021)

Publications as first author released within the framework of this cumulative dissertation:

1. **Wissner, J. L.**, Escobedo-Hinojosa, W., Heinemann, P. M., Hunold, A. & Hauer, B. Methods for the detection and analysis of dioxygenase catalyzed dihydroxylation in mutant derived libraries. in *Methods in Enzymology*, **644**, 63-93 (2020). doi:10.1016/bs.mie.2020.04.022

2. **Wissner, J. L.**, Ludwig, J., Escobedo-Hinojosa, W. & Hauer, B. An enhanced toluene dioxygenase platform for the production of cis-1,2-dihydrocatechol in Escherichia coli BW25113 lacking glycerol dehydrogenase activity. *Journal of Biotechnology*, **325**, 380-388 (2020). doi:10.1016/j.jbiotec.2020.09.012

3. **Wissner, J. L.**, Ludwig, J., Escobedo-Hinojosa, W. & Hauer, B. Strategy for identification of cis-dihydrodiendiol-degrading dehydrogenases in E. coli BW25113. *MethodsX*, **7**, 101143 (2020). doi:10.1016/j.mex.2020.101143

4. **Wissner, J. L.**, Escobedo-Hinojosa, W., Vogel, A. & Hauer, B. An engineered toluene dioxygenase for a single step biocatalytical production of (-)-(1S,2R)-cis-1,2-dihydro-1,2-naphthalenediol. *Journal of Biotechnology*, **326**, 37-39 (2021). doi:10.1016/j.jbiotec.2020.12.007

5. **Wissner, J . L.**, Schelle, J. T., Escobedo-Hinojosa, W., Vogel, A. & Hauer, B. Semi-Rational Engineering of Toluene Dioxygenase from Pseudomonas putida F1 towards Oxyfunctionalization of Bicyclic Aromatics. *Advanced Synthesis & Catalysis*, **363**, (2021). doi:10.1002/adsc.202100296

6. **Wissner, J . L.**, Schelle, J. T., Escobedo-Hinojosa, W., Vogel, A. & Hauer, B. *Front Cover Picture*. Semi-Rational Engineering of Toluene Dioxygenase from Pseudomonas putida F1 towards Oxyfunctionalization of Bicyclic Aromatics. *Advanced Synthesis & Catalysis*, **363**, (2021). doi:10.1002/adsc.202100532

The following publications were coauthored and are not part of this cumulative dissertation:

7. Escobedo-Hinojosa, W., **Wissner, J. L.**, Hauer, B. A real-time 31P-NMR-based approach for the assessment of glycerol kinase catalyzed monophosphorylations. *MethodsX*, **8**, 101285 (2021). doi:10.1016/j.mex.2021.101285

8. Farr, T., **Wissner, J. L.**, Hauer, B. A simple and efficient method for lyophilization of recombinant E. coli JM109 (DE3) whole-cells harboring active Rieske non-heme iron dioxygenases. *MethodsX*, **8**, 101323 (2021). doi:10.1016/j.mex.2021.101323

CONTRIBUTION REPORT

1. Responsible for writing the manuscript. The coauthors revised the first version of the manuscript.

2. Responsible for writing the manuscript and for all experimental work. Characterization of various reaction and expression condition was performed by Julian Ludwig in the course of his master thesis. The coauthors revised the first version of the manuscript.

3. Responsible for writing the manuscript and for all experimental work. The coauthors revised the first version of the manuscript.

4. Responsible for writing the manuscript and for all experimental work. Variants were generated by c-LEcta, Leipzig (DE). The coauthors revised the first version of the manuscript.

5. Responsible for writing the manuscript and for all experimental work. Screening of 1,2,3,4-tetrahydroquinoline with the single point mutant library, was performed by Jona T. Schelle in the course of his master thesis. Variants were generated by the c-LEcta, Leipzig (DE). The coauthors revised the first version of the manuscript.

6. Responsible for the creative conception of the image. The picture was designed by *WesWizzArt* and painted by the Mexican artist Way (Wayra Castillo-Guerrero).

ACKNOWLEDGEMENTS

ABSTRACT

Till today, the selective dearomatizing *cis*-dihydroxylation of aromatics cannot be performed chemically. Hence, it is truly exceptional how nature evolved Rieske non-heme iron dioxygenases (ROs), capable of performing such challenging reactions. These enzymatic multicomponent systems are not only able to catalyze *cis*-dihydroxylation reactions with outstanding activity, but also with an excellent enantioselectivity (>99% *ee*) for a broad range of aromatic substrates. In their natural host, ROs are involved in the initial catabolism of aromatic compounds into intermediates of the tricarboxylic acids cycle. In biocatalysis, the oxyfunctionalizing capabilities of ROs can be utilized to synthetize *cis*-dihydrodiols from mono- and polycyclic arenes. These hydroxylated compounds are valuable synthons, frequently employed in the synthesis of natural products, pharmaceuticals, and fine chemicals. One of the best experimentally characterized ROs reported in literature is toluene dioxygenase (TDO), from *Pseudomonas putida* F1, which substrate scope comprises over 100 different compounds. Nevertheless, a restraint hampering TDO application as a biocatalyst for the targeted *cis*-dihydroxylation of aromatics, is the observed decrease in conversion as substrate size increases, resulting in modest product yields. Therefore, fundamental studies revealing the influence of TDO active-site positions, enabling a better conversion of mono- and especially bulky bicyclic aromatics are compulsory. Such need prompted the present thesis, which focused on TDO as model biocatalysts, since it displays remarkable substrate promiscuity and advantageous features for preparative biotransformations. The central task of this work was fostering the biocatalytic capabilities of TDO toward the *cis*-dihydroxylation of four aromatic model substrates. The first step was the development of an enhanced recombinant TDO system in *Escherichia coli* (*E. coli*), for an efficient dioxygenase overexpression. Therefore, the new platform *E. coli* BW25113 pBAD18-TDO was established, which exhibited outstanding product formation for the bicyclic substrate naphthalene. Nevertheless, in the attempt to apply such platform for the conversion of small monocyclic aromatics, an unforeseen downstream reaction catalyzed by *E. coli*, dehydrogenating the generated *cis*-dihydrocatechols to the corresponding catechols, was discovered. By performing a systematic screening of dehydrogenase deficient single knock-out strains from the KEIO collection, the enzyme glycerol dehydrogenase (GldA) from *E. coli* was identified as the main responsible for such degradation. These findings drove the development of the enhanced platform *E. coli* BW25113 Δ*gldA* pBAD18-TDO, allowing the abolishment of the unwanted secondary reaction. Thus, a semi-preparative biotransformation of benzene was performed utilizing the customized TDO

platform, resulting in the isolation of 141 mg (31%) of *cis*-dihydrocatechol as sole product. The system was also tested for the semi-preparative biotransformation of the bicyclic substrate naphthalene, yielding 287 mg (89%) of enantiopure (1*R*,2*S*)-1,2-dihydro-1,2-naphthalenediol. Once the TDO platform was established and successfully applied, the next task was to explore the influence of TDO active-site position F366 in naphthalene conversion by generating and investigating a set of nine TDO variants at this position. Strikingly, the single point variant TDO$_{F366V}$ revealed that the enantioselectivity could be switched completely. Furthermore, semi-preparative naphthalene biotransformations with TDO$_{F366V}$ enabled the synthesis of 101 mg (31%) enantioenriched (1*S*,2*R*)-1,2-dihydro-1,2-naphthalenediol (90% *ee*). It is worth to mention that before this study, this enantiomer was never directly generated either chemically or biocatalytically. The next aim was to expand the TDO-based biocatalyst portfolio by generating a semi-rational designed TDO single- and double mutant library. Thus, the library consisting out of 176 variants was tested for the conversion of the bicyclic substrates naphthalene, 1,2,3,4-tetrahydroquinoline, and 2-phenylpyridine in the pursue to enhance product formation and/or chemo-, regio- and enantioselectivity. These studies highlighted that introduced mutations at the active site hot-spot positions M220, A223 and F366, strongly influences chemo-, regio- and enantioselectivity. Moreover, mutations at positions M220 and A223 also exerted substantial effects on product formation during the conversion of bicyclic (hetero)aromatics. In addition, since the TDO mutant library addressed all 14 non-conserved active site amino acids, it was noticeable that the active site is highly tolerant to the introduction of mutations. Additionally, it could be assessed that the combination of the outperforming mutations into the double variant TDO$_{F114H\_A223T}$ entirely abolished the formation of the side product quinoline in 1,2,3,4-tetrahydroquinoline biotransformations. This approach enabled in a semi-preparative biotransformation the selective production of 106 mg (71%) of enantioenriched (*R*)-1,2,3,4-tetrahydroquinoline-4-ol (94% *ee*). Further, double variant TDO$_{M220A\_V309G}$ exhibited an astonishing 15.1-fold higher conversion of the substrate 2-phenyl-pyridine, in comparison to TDO wild type. This enhancement enabled for the first time, the TDO-catalyzed production of 114 mg (60%) enantiopure (1*S*,2*R*)-3-(pyridin-2-yl)cyclohexa-3,5-diene-1,2-diol, along with reduced amounts of 6 mg (4%) 2-phenylpyridin-3-ol, as side product. These compelling findings meet the scope of this thesis, in terms of fostering TDO product formation as well as chemo-, regio- and enantioselectivity for bulky bicyclic (hetero) aromatics. Hence, this dissertation highlights the importance of both, the improvement of the recombinant *E. coli* BW25113 platform, as well as the enhancement of the biocatalyst TDO *via* enzyme engineering for the generation of valuable *cis*-dihydroxylated compounds.

ZUSAMMENFASSUNG

Bis zum heutigen Tag kann die selektive dearomatisierende *cis*-Dihydroxylierung von Aromaten nicht chemisch durchgeführt werden. Daher ist es wirklich außergewöhnlich, wie die Natur Rieske nicht-Häm Eisen Dioxygenasen (ROs) entwickelt hat, die in der Lage sind, solch anspruchsvolle Reaktionen durchzuführen. Diese enzymatischen Multikomponentensysteme sind nicht nur in der Lage, *cis*-Dihydroxylierungsreaktionen mit hervorragender Aktivität zu katalysieren, sondern auch mit exzellenten Enantioselektivität (>99% *ee*) für eine breite Auswahl an aromatischen Substraten. In ihrem natürlichen Wirt sind ROs am initialen Katabolismus von aromatischen Verbindungen zu Intermediaten des Tricarbonsäurezyklus beteiligt. In der Biokatalyse können die oxyfunktionalisierenden Eigenschaften der ROs genutzt werden, um *cis*-Dihydrodiole aus mono- und polyzyklischen Arenen zu synthetisieren. Diese hydroxylierten Verbindungen sind wertvolle Synthone, die häufig in der Synthese von Naturstoffen, Pharmazeutika und Feinchemikalien eingesetzt werden. Eine der am besten experimentell charakterisierten ROs, über die in der Literatur berichtet wird, ist die Toluoldioxygenase (TDO) aus *Pseudomonas putida* F1, deren Substratspektrum über 100 verschiedene Verbindungen umfasst. Ein Hemmnis für den Einsatz der TDO als Biokatalysator für die gezielte *cis*-Dihydroxylierung von Aromaten ist jedoch die beobachtete Abnahme des Umsatzes mit zunehmender Substratgröße, was zu mäßigen Produktausbeuten führt. Daher sind grundlegende Untersuchungen zum Einfluss der TDO-Positionen im aktivem Zentrum, die eine bessere Umsetzung von mono- und insbesondere sperrigen bizyklischen Aromaten ermöglichen, zwingend erforderlich. Diese Notwendigkeit regte die vorliegende Thesis an, welche sich auf TDO als Modellbiokatalysator konzentrierte, da dieser eine bemerkenswerte Substratpromiskuität und vorteilhafte Eigenschaften für präparative Biotransformationen aufweist. Die zentrale Aufgabe dieser Arbeit war die Förderung der biokatalytischen Potentials der TDO zur *cis*-Dihydroxylierung von vier aromatischen Modellsubstraten. Der erste Schritt war die Entwicklung eines verbesserten rekombinanten TDO-Systems in *Escherichia coli* (*E. coli*), für eine effiziente Überexpression der Dioxygenase. Dazu wurde die neue Plattform *E. coli* BW25113 pBAD18-TDO etabliert, welche eine hervorragende Produktbildung für das bizyklische Substrat Naphthalin zeigte. Jedoch wurde bei dem Versuch, diese Plattform für die Umsetzung kleiner monozyklischer Aromaten einzusetzen, eine unvorhergesehene, von *E. coli* katalysierte nachgeschaltete Reaktion entdeckt, nämlich die Dehydrierung der erzeugten *cis*-Dihydrocatechole zu den entsprechenden Catecholen. Durch ein systematisches Screening von dehydrogenasedefizienten Einzelknockout-Stämmen aus der KEIO-Sammlung wurde das

Enzym Glycerindehydrogenase (GldA) aus *E. coli* als Hauptverantwortlicher für diesen Abbau identifiziert. Diese Ergebnisse trieben die Entwicklung der verbesserten Plattform *E. coli* BW25113 Δ*gldA* pBAD18-TDO voran, die die Beseitigung der unerwünschten Nebenreaktion ermöglichte. So wurde unter Verwendung der maßgefertigten TDO-Plattform eine semipräparative Biotransformation mit Benzol durchgeführt, was zur Isolierung von 141 mg (31%) *cis*-Dihydrocatechol als einziges Produkt führte. Das System wurde auch für die semipräparative Biotransformation des bizyklischen Substrats Naphthalin getestet und lieferte 287 mg (89%) enantiomerenreines (1*R*,2*S*)-1,2-Dihydro-1,2-naphthalindiol. Nachdem die TDO-Plattform etabliert und erfolgreich angewendet wurde, bestand die nächste Aufgabe darin, den Einfluss der Position F366 im aktiven Zentrum der TDO bei der Naphthalinumsetzung zu erkunden, indem ein Satz von neun TDO-Varianten an dieser Position erzeugt und untersucht wurde. Bemerkenswerterweise zeigte die Einzelpunktvariante TDO$_{F366V}$, dass die Enantioselektivität vollständig umgekehrt werden konnte. Darüber hinaus ermöglichte die semipräparative Naphthalin-Biotransformation mit TDO$_{F366V}$ die Synthese von 101 mg (31%) enantiomerenangereichertem (1*S*,2*R*)-1,2-Dihydro-1,2-naphthalindiol (90% *ee*). Es ist erwähnenswert, dass dieses Enantiomer vor dieser Studie weder chemisch noch biokatalytisch direkt hergestellt wurde. Das nächste Ziel war die Erweiterung des TDO-basierten Biokatalysatorbestandes durch die Generierung einer semi-rational konstruierten TDO-Einzel- und Doppelmutanten-Bibliothek. So wurde die aus 176 Varianten bestehende Bibliothek für die Umsetzung der bizyklischen Substrate Naphthalin, 1,2,3,4-Tetrahydrochinolin und 2-Phenylpyridin mit dem Ziel getestet, die Produktbildung und/oder die Chemo-, Regio- und Enantioselektivität zu verbessern. Diese Studien hoben hervor, dass eingeführte Mutationen an den Hot-Spot-Positionen des aktiven Zentrums, M220, A223 und F366, die Chemo-, Regio- und Enantioselektivität stark beeinflussen. Darüber hinaus hatten die Mutationen an den Positionen M220 und A223 auch beträchtliche Auswirkungen auf die Produktbildung bei der Umsetzung von bizyklischen (Hetero-)Aromaten. Da die TDO-Mutantenbibliothek alle 14 nicht konservierten Aminosäuren des aktiven Zentrums adressierte, war außerdem zu erkennen, dass das aktive Zentrum sehr tolerant gegenüber der Einführung von Mutationen ist. Zusätzlich konnte festgestellt werden, dass die Kombination der überragenden Mutationen zur Doppelvariante TDO$_{F114H\_A223T}$ bei Biotransformationen von 1,2,3,4-Tetrahydrochinolin die Bildung des Nebenprodukts Chinolin vollständig beseitigte. Dieser Ansatz ermöglichte in einer semipräparativen Biotransformation die selektive Herstellung von 106 mg (71%) enantiomerenangereichertem (*R*)-1,2,3,4-Tetrahydrochinolin-4-ol (94% *ee*). Darüber hinaus zeigte die Doppelvariante TDO$_{M220A\_V309G}$, im Vergleich zum TDO-Wildtyp, eine erstaunliche

15,1-fach höhere Umsetzung des Substrats 2-Phenylpyridin. Diese Verbesserung ermöglichte erstmals die TDO-katalysierte Produktion von 114 mg (60%) enantiomerenreinem (1*S*,2*R*)-3-(Pyridin-2-yl)cyclohexa-3,5-dien-1,2-diol, zusammen mit reduzierten Mengen an 2-Phenylpyridin-3-ol 6 mg (4%), als Nebenprodukt. Diese überzeugenden Ergebnisse entsprechen dem Ziel dieser Arbeit, die TDO-Produktbildung sowie die Chemo-, Regio- und Enantioselektivität für sperrige bizyklische (Hetero-)Aromaten zu fördern. Diese Dissertation unterstreicht daher die Bedeutung sowohl der Verbesserung der rekombinanten *E. coli* BW25113-Plattform als auch der Optimierung des Biokatalysators TDO durch Enzym-Engineering für die Generierung von wertvollen *cis*-dihydroxylierten Verbindungen.

ABBREVIATIONS

Å	Å [10^{-10} m]
AD	asymmetric dihydroxylation
BPDO	biphenyl dioxygenase from *Burkholderia xenovorans* LB400 or *Pseudomonas pseudoalcaligenes* KF707
CDO	cumene dioxygenase from *Pseduomoas fluorescens* IP01
cww	cell wet weight
DAD	diode-array detector
DHQ	dihydroquinine
DHQD	dihydroquinidine
DNA	deoxyribonucleic acid
E. coli	*Escherichia coli*
ee	enantiomeric excess [%]
EntA	2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase from *E. coli*
ESI-MS	electrospray ionization mass spectrometer
GldA	glycerol dehydrogenase from *Escherichia coli* BW25113
HcaB	3-phenylpropionate-dihydrodiol dehydrogenase from *E. coli*
HPLC	high-pressure liquid chromatography
IND	indoline-like ligand class
IPTG	isopropyl β-D-1-thiogalactopyranoside
M	molar concentration [mol L^{-1}]
NAD(P)H	nicotinamide adenine dinucleotide (phosphate)
NDO	naphthalene dioxygenase from *Pseudomonas species* NCIB 9816-4
NMO	*N*-methylmorpholine *N*-oxide
PCR	polymerase chain reaction
PDB	protein database
pH	power of hydrogen (negative base 10 logarithm of the H$^+$ ion activity)
PHAL	phthalazine-like ligand class
P. putida	*Pseudomonas putida*
PYR	pyrimidine-like ligand class
RO	Rieske non-heme iron dioxygenase
TEC	chlorobenzene dioxygenase from *Burkholderia species* strain PS12
TDO	toluene dioxygenase from *Pseudomonas putida* F1
TodA	flavoprotein reductase of the TDO multicomponent system
TodB	ferredoxin of the TDO multicomponent system
TodC1	oxygenase α-subunit of the TDO multicomponent system
TodC2	oxygenase β-subunit of the TDO multicomponent system
TodD	*cis*-toluene dihydrodiol dehydrogenase of *P. putida* F1

TABLE OF CONTENTS

1 INTRODUCTION

1.1 Biocatalysis

Over the past decades, the application of enzymes as biocatalysts have proved to be an efficient and attractive alternative to chemical catalysts, becoming especially indispensable in pharmaceutical, specialty chemicals, and flavor and fragrance industry.[1,2] Biocatalysts are green alternatives to the toxic chemical heavy metal catalysts and exhibit in addition often outstanding chemo-, regio- and enantioselectivity.[3] Moreover, these environmentally friendly catalysts work under mild reaction conditions, employing water as the sole solvent and displaying a remarkable performance at room temperature. Furthermore, the use of enzymes and/or whole cells as biocatalysts in organic synthesis is often leading to the circumvention of multiple reaction steps, contributing to more economic industrial processes.[4] Actually, over the last decades several chemical and pharmaceutical companies, such as BASF (DE), DSM (NL), Codexis (US), DuPont (US), Merck & Co. (US) and Pfizer (US), have begun to significantly increase the number of biocatalytic processes used in their synthesis operations.[5,6] Nevertheless, there are still some challenges often hindering the substitution of established chemical processes by biocatalytic ones. For instance, lack of substitute enzymes for multiple valuable chemical reactions, since such biocatalysts have either not yet been discovered or simply not been evolved by nature. In addition, existing enzymes can exhibit low product formation or a different selectivity than for the targeted product. In this sense, enzyme engineering *via* mutagenesis is a powerful tool and an attractive solution substantially revolutionizing biocatalysis. Limitations such as reaction spectrum, enzyme stability, selectivity and low activity can be overcome by utilizing various mutagenesis strategies. The introduction of mutations is mostly based on the polymerase chain reaction (PCR), since this technique is easy to perform and capable of creating extensive genetic diversity.[7] Routine methods for the generation of mutant libraries include DNA shuffling, error-prone PCR (epPCR) and site-saturation mutagenesis.[8] The state-of-the-art in mutation strategies currently applied, can be divided into four major categories;

1) Directed evolution *via* epPCR, gene-shuffling or site saturation mutagenesis.[9,10] This approach generates improved mutants by means of random mutagenesis, subsequent selection, and identification. Thus, random approaches attempt to mimic Darwinian evolution with the advantage that no structural information of the enzyme is needed and only the protein sequence is required. However, usually large mutant libraries have to

be screened to isolate variants with beneficial characteristics. Since the relevant impact and scope of directed evolution of enzymes, in 2018 Frances Arnold was laureated with the Nobel Prize in Chemistry for her pioneering work in this field.[11]

2) Rational enzyme design *via* site-directed mutagenesis.[12,13] This strategy allows the substitution of specific amino acids into the enzyme of interest, aiming to alter its physical or chemical properties. Therefore, previous knowledge of the catalytic site revealed by the crystal structure of the enzyme is crucial for this approach, to achieve improved catalytic activity, convert unnatural substrates, or enable the catalysis of new reactions.

3) Semi-rational enzyme design.[14,15] The combination of evolution and rational approaches is commonly known as semi-rational design, where small targeted libraries are generated. The generation of these smart libraries is based on structural or catalytic knowledge, allowing multiple, specific residues to be mutated, leading to positive results at a higher success rate.

4) *De novo* enzyme design.[16–18] Here, computational enzyme design is utilized to create enzymes from scratch, catalyzing novel reactions. The purpose is to simultaneously introduce both, the substrate binding pocket and the catalytic machinery into protein scaffolds devoid of the desired activity.

Thanks to the advances in enzyme engineering over the last three decades, an increasingly broad spectrum of bioreactions is available. Such uprising of biocatalysis has emerged to provide alternatives for the synthesis of a broad variety of sough-after molecules, including many valuable *cis*-dihydroxylated synthons. The latter point encouraged the investigations of the present doctoral thesis.

1.2 Relevance of *cis*-dihydroxylated synthons
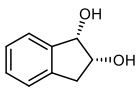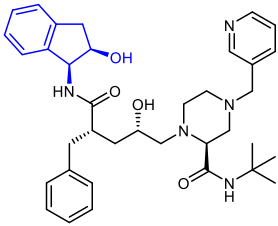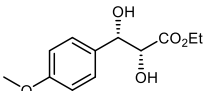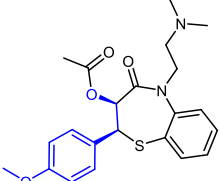
Olefins as well as aromatics are readily available and versatile starting materials, which are commonly applied in organic chemistry for the generation of useful compounds harboring pursued biological activities.[19,20] Hence, one of the key foci in organic synthesis is the implementation of chirality into prochiral C=C double bonds.[21] In this manner, the regio- and enantioselective *cis*-dihydroxylation of non-activated olefins and aromatic compounds, performed either chemically or biocatalytically, is of high synthetic value.[22–26] Indeed, optically active *cis*-diols are the most sought-after compounds in asymmetric synthesis, by virtue of the effortless substitution of the diols into several other functional groups.[27,28] In this

Introduction

manner, *cis*-diols are frequently utilized as building blocks in the chemical and pharmaceutical industry (Table 1).[1,29–32]

One notable archetype compound is *cis*-(1*S*,2*R*)-indandiol **1**, a key intermediate in the synthesis of the protease inhibitor indinavir **2** (Crixivan®), applied for AIDS treatment.[29] Another relevant example is the synthon ethyl (2*S*,3*R*)-2,3-dihydroxy-3-(4-methoxyphenyl)propionate **3**, which can be further used for the generation of the heart drug diltiazem **4**.[33] Besides, the synthon (1*R*,2*S*)-1,2-dihydro-1,2-naphthalenediol **5a** is used in the synthesis of (+)-goniodiol **6**, which exhibits antitumoral activity.[34] Furthermore, the *cis*-dihydroxylated molecule benzyl (2*R*,3*R*,4*S*)-3,4-dihydroxy-5-(4-methoxyphenyl)-5-oxo-pentan-2-ylcarbamate **7** finds application in the total synthesis of the antibiotic compound (-)-codonopsinine **8**.[35] Lately, Boyd and colleagues successfully employed the *cis*-dihydroxylated synthon (1*R*,2*S*)-1,2-dihydrodibenzo[*b,d*]furan-1,2-diol **9**, resulting in a shorter synthetic sequence for the generation of the fungal metabolite (-)-ribisin A **10** with interesting pharmaceutical applicability.[36] The chemical space of *cis*-dihydroxylated synthons is certainly very broad as has been illustrated by Hudlicky and Heravi and colleagues, reporting an extensive repertoire of relevant examples for the utilization of *cis*-dihydroxylated compounds in the synthesis of natural products.[24,32]

Table 1: Selected examples of pharmaceutically relevant bioactive compounds generated from *cis*-dihydroxylated scaffolds. The backbones of the dihydroxylated synthons are depicted in blue in the illustrated compounds.

cis-dihydroxylated synthon	Pharmaceutical relevant product	Literature reference
1	**2**	[29]
3	**4**	[33]
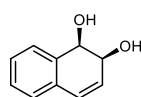
5a
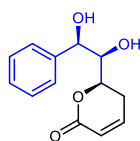
6

[34]
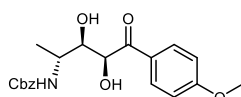
7
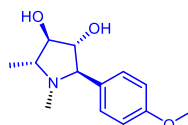
8

[35]
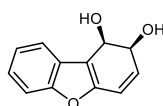
9
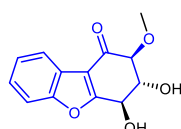
10

[36]

1.3 Chemically driven *cis*-dihydroxylations

Generally, the chemical synthesis of *cis*-dihydroxylated compounds utilizes olefins as substrates, incorporating selectively two hydroxyl groups in one C=C double bond *via* heavy metal-catalyzed asymmetric synthesis.[37] Actually, the current industrial production of *cis*-dihydroxylated compounds is based on such approach.[38] Thereby, the direct oxidative functionalization of C=C double bonds is catalyzed by heavy metal oxides, e.g. of the transition metals osmium, ruthenium or manganese.[22,39] In addition, *cis*-dihydroxylations mediated by iron, molybdenum, palladium, cerium, technetium, selenium, oxone, and hypervalent iodine are reported in literature. However such approaches are uncommonly employed with exception of the former.[22,40] Hence, the majority of published results within the last decades are based on the heavy metal catalyst osmium tetroxide (OsO$_4$).[41]

1.3.1 Osmium mediated cis-*dihydroxylation*

The ability of osmium tetroxide (OsO$_4$) to oxidize olefins has been known now for over a century.[42] In 1913, Hoffmann showed that OsO$_4$ could be regenerated in the presence of sodium or potassium chlorate for the hydroxylation of olefins, enabling a catalytic process.[43] However, the applicability of this reaction was hampered by the observed overoxidation of the products, resulting in separation problems and decreased yields. Later on, Criegee and colleagues showed that OsO$_4$ could be used effectively for the *cis*-dihydroxylation of various

olefins, when used stoichiometrically without a cooxidant.[44,45] Furthermore, Criegee suggested that the reaction takes place *via* the formation of an Os(VI) ester complex, which is hydrolyzed reductively to yield *cis*-hydroxylated diols (Figure 1).
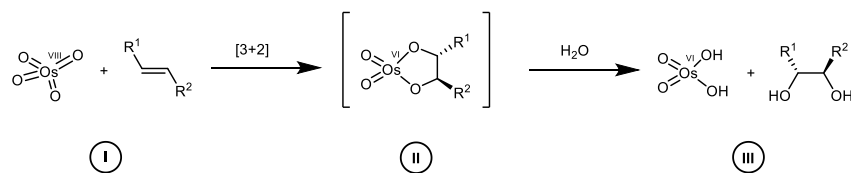
Figure 1: Stoichiometric osmium tetroxide-catalyzed *cis*-dihydroxylation of an olefin. The olefin (**I**) along with osmium tetroxide (OsO$_4$) forms a cyclic Os(VI) intermediate (**II**), which undergoes a [3+2]-cycloaddition reaction and further reductive hydrolysis, yielding the corresponding *cis*-diol and osmium hydroxide oxide (Os(OH)$_2$O$_2$) (**III**). Adapted from Döbler and colleagues.[46]

In 1976, VanRheenen and colleagues developed the Upjohn dihydroxylation, utilizing OsO$_4$ as catalyst.[47] The peculiarity of this organic reaction is the use of *N*-methyl-morpholine *N*-oxide (NMO), which acts as cooxidant for OsO$_4$ regeneration. This resulted in a superior catalytic process, to the previously reported stoichiometric OsO$_4$ syntheses. Drawbacks of this reaction are the formation of a racemic *cis*-diol product mixtures, the liability to over oxidation, and the incapability to convert tetrasubstituted olefins.[48] Later on, in 1988 Sharpless and collaborators developed the Sharpless asymmetric dihydroxylation (AD),[49] which was based on the Upjohn dihydroxylation. Here, olefins are *cis*-dihydroxylated by OsO$_4$ in the presence of derivatives of the chiral cinchona alkaloids dihydroquinine (DHQ) or dihydroquinidine (DHQD), which perform as ligands, and the cooxidant NMO. Thereby, the cinchona alkaloid derivatives are coordinated to the osmium catalyst, forming a chiral complex resembling an enzyme-like active pocket. This chiral osmium-ligand complex can distinguish between the prochiral faces of olefin substrates, leading to the enantioenriched *cis*-diol formation in high yields. Furthermore, regarding the chosen ligand, different *cis*-diol enantiomers were formed for the majority of the tested substrates. However, the enantiomeric excess of the *cis*-diol products obtained were lower than those produced by the stoichiometric reaction. The explanation for this discrepancy could be due to the occurrence of a second catalytic cycle with a lower enantio-selectivity (Figure 2).[50]
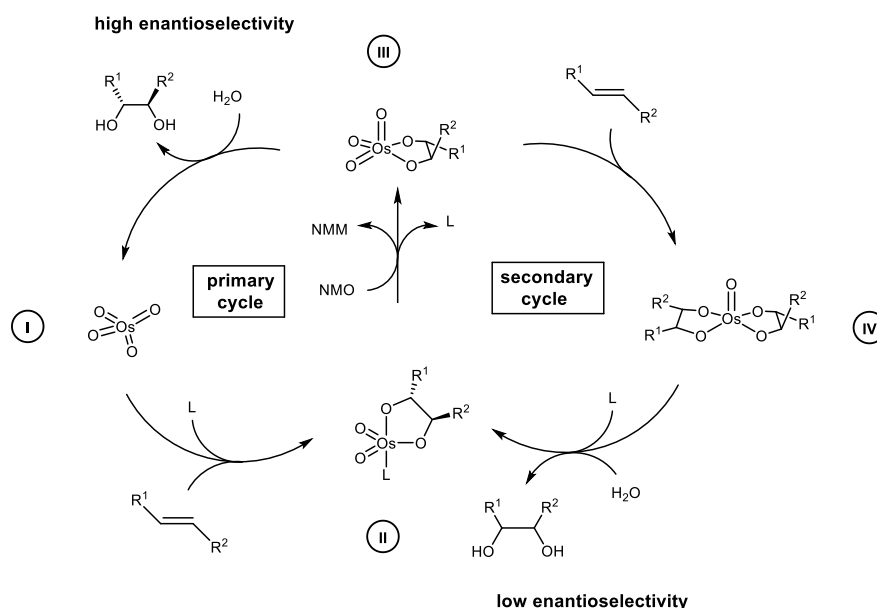
Figure 2: The two co-occurring catalytic cycles of the asymmetric dihydroxylation using NMO as cooxidant in an aqueous one phase system. OsO_4 (**I**) with the cinchona alkaloid ligand (L) forms an OsO_4 – ligand complex, which undergoes a [3+2]-cycloaddition with the olefin substrate, to yield the cyclic intermediate (**II**). Intermediate (**II**) is oxidized by the cooxidant *N*-methylmorpholine *N*-oxide (NMO), forming trioxo osmium(VIII) glycolate (**III**), *N*-methylmorpholine (NMM), and releasing the ligand (L). Hydrolysis of compound (**III**) yields osmium tetroxide (**I**) and the *cis*-dihydroxylated product with high enantiomeric excess (primary cycle). Alternatively, compound (**III**) can coordinate another olefin substrate to yield compound (**IV**). Hydrolysis and addition of one ligand (L) yields the intermediate (**II**) and the *cis*-dihydroxylated product with low enantiomeric excess (secondary cycle). Adapted from Kolb and coworkers.[27]

Thus, the counterproductive secondary cycle exhibits in comparison to the primary cycle, not only lower enantiomeric excesses for the dihydroxylated products, but also reduced turnover frequencies. This prompted Sharpless and coworkers to develop a biphasic system with potassium ferricyanide ($K_3Fe(CN)_6$) as cooxidant, which successfully abolished the secondary cycle completely (Figure 3).[51] Thereby, the osmylation of the olefin substrate proceeds in the organic phase and the regeneration of the Os(VI) species in the aqueous phase. The catalytic $K_3Fe(CN)_6$ system achieves high enantioselectivities identical to those obtained in stoichiometric AD.
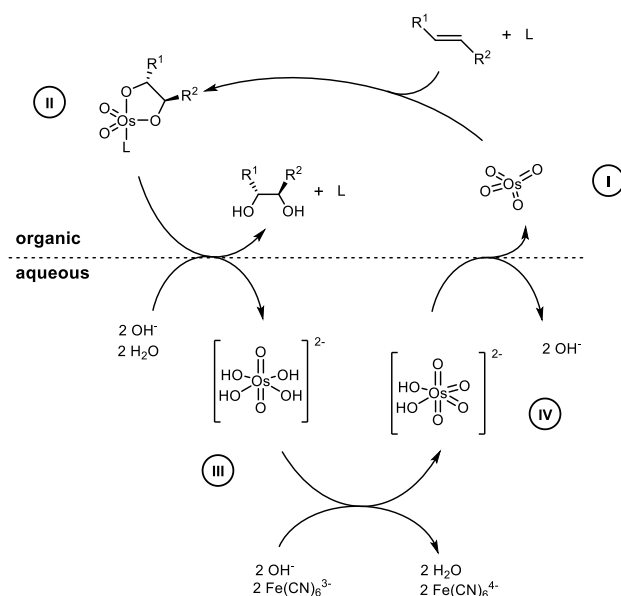
Figure 3: Catalytic cycle of the AD reaction with $K_3Fe(CN)_6$ as cooxidant in a two phase system. OsO_4 (**I**) along with the cinchona alkaloid ligand (L) forms in the organic phase an OsO_4 – ligand complex, which undergoes a [3+2]-cycloaddition with the olefin substrate, to yield the cyclic intermediate (**II**). Intermediate (**II**) then undergoes hydrolysis, releasing the *cis*-hydroxylated diol and the ligand back to the organic phase, while the resulting Os(VI) species migrate into the aqueous phase as $OsO_2(OH)_4^{2-}$ (**III**). The oxidation of (**III**) by ferricyanide ($Fe(CN)_6^{3-}$), along with the loss of two H_2O molecules drives the $OsO_4(OH)_2^{2-}$ species (**IV**) toward the regeneration of OsO_4 (**I**), which migrates back to the organic phase. Adapted from Ogino and colleagues.[51]

Despite the exploration of more than 400 alkaloid ligands during the last decades, the cinchona alkaloid derivatives of DHQ and DHQD remain the most effective for the Sharpless AD reaction.[52] It was shown, that the enantioselectivity can be influenced by residues, attached to the cinchona alkaloid backbone. Thereby, the three residue classes phthalazine-like (PHAL), pyrimidine-like (PYR) and indoline-like (IND) outperformed for multiple substrates, enabling the *cis*-dihydroxylation of almost any olefin (Figure 4).[27,41]

Introduction
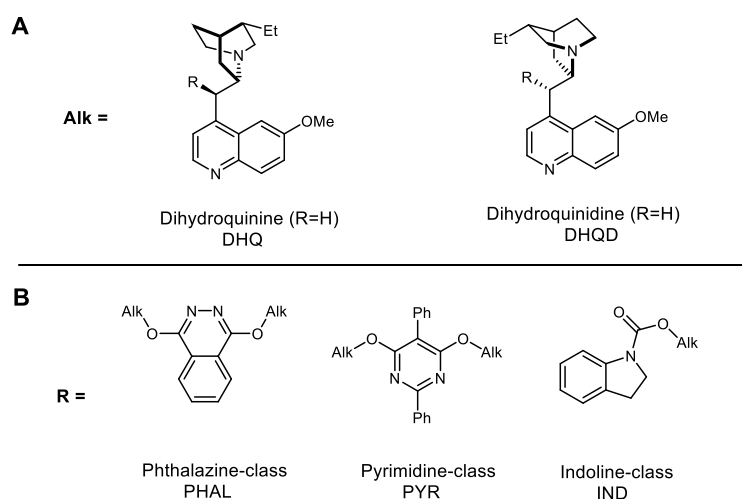
Figure 4: Cinchona alkaloid derivatives used as ligands in Sharpless AD. (**A**) Backbone of the cinchona alkaloid (Alk) diastereomers dihydroquinine (DHQ) and dihydroquinidine (DHQD). (**B**) Replacement of the cinchona alkaloid residues (R) with phthalazine-like (PHAL), pyrimidine-like (PYR), and indoline-like (IND) structures spawns the most versatile and effective ligands for Sharpless AD. Adapted from Chang and coworkers.[52]

In addition, these three different cinchona-based ligands effectively enabled the *cis*-dihydroxylation of all six possible olefin substitution patterns, namely mono-, *gem*-di, *cis*-di, *trans*-di, tri, and tetrasubstituted olefins with outstanding enantioselectivity (up to 99.8% *ee*).[27,53] However, even by applying the special IND ligand, the enantioselective Sharpless AD of *cis*-disubstituted olefins remains a challenge, exhibiting only moderate enantiomeric excess (*ee*) values < 80%. Though, by applying DHQ- and DHQD-derived ligands with the corresponding PYR, PHAL, or IND residues, both enantiomers of a *cis*-diol are accessible. These improvements, proved the Sharpless AD to be one of the most attractive procedures to produce enantiomerically pure *cis*-diols, during the last decades.[27] Furthermore, it is the method of choice for the generation of *cis*-hydroxylated diols, owed to its high yields, outstanding enantioselectivity, broad substrate scope, as well as its high tolerance towards a variety of functional groups at the substrate molecule.[54–56] Thus, to date, the OsO$_4$ based AD finds numerous applications for the total synthesis of natural products.[32] In addition, it is noteworthy to mention, that Karl B. Sharpless was laureated in 2001 with the Nobel prize in Chemistry,[57] for his work on chirally catalyzed oxidation reactions enabled by the Sharpless epoxidation[58] and AD.[49] Hence, this underlines not only the importance and the extend applicability of AD, but also the paramountcy of the *cis*-dihydroxylated compounds. However, one of the main disadvantages of the Sharpless AD is that the *cis*-dihydroxylation of aromatics is not within the scope of this reaction. Further drawbacks are the high toxicity of osmium compounds, their high costs, as well as their volatility, and the negative environmental impact due to organic solvents and heavy metal catalysts applied.[22,54]

1.3.2   Ruthenium, manganese and iron mediated cis-*dihydroxylations*

In reaction to the drawbacks of OsO4, several research groups have investigated alternative reagents for the *cis*-dihydroxylation of olefins.[22] Therefore, one of the most common catalyst used is the isoelectronic ruthenium tetroxide (RuO4), which in comparison to OsO4 is less toxic and cheaper. The oxidative RuO4 catalysis has been shown to *cis*-dihydroxylate a wide range of olefins, with exception of strained bicyclic olefins, sterically hindered tri- and most tetrasubstituted olefins.[59] In 1994, Shing and colleagues described the rapid *cis*-dihydroxylation of a variety of olefins, utilizing RuCl3·H2O as catalyst and NaIO4 as cooxidant.[60] Strikingly, the products were obtained in less than three minutes and in good yields. Plietker and colleagues further performed the enhancement of the reaction and mechanistic studies, resulting in a decrease in catalyst loading and in the conversion of a larger array of olefins in excellent yields.[61–63] Nevertheless, under the applied reaction conditions the RuO4 catalyzed generation of *cis*-dihydroxylated compounds is not enantioselective, thus a racemic product mixture is obtained.[59] To date, only a small number of publications are reporting the enantioselective RuO4-catalyzed AD. One example is the enantioselective cross-methathesis-dihydroxylation of olefins, using chiral auxiliary substituted acrylamides.[64] In such manner, the Grubbs-Hoyveda methathesis catalyst is converted *in situ* into RuO4 to perform the *cis*-dihydroxylation step. This approach represents the first efficient diastereoselective RuO4-catalyzed oxidation, enabling good yields and outstanding enantioselectivities (90-99% *ee*). However, the major drawback of ruthenium based catalysts, hindering their widespread use, is the high oxidative potential of Ru(VIII) species, leading often to overoxidation.[22] The different reactivity between OsO4 and RuO4 can be explained by the higher stability of Os(VIII) compared to Ru(VIII) complexes.[65] Thus, RuO4 is more commonly applied in ketohydroxylation and oxidative cleavage reactions, than for *cis*-dihydroxylation of olefins.[59,63] An alternative catalyst for the generation of *cis*-diol from olefins, is permanganate (MnO4$^-$), isoelectronic to OsO4 and RuO4. One of the first reported manganese catalyzed *cis*-dihydroxylation is the process developed by De Vos and colleagues in 1999.[66] They described that substantial amounts of *cis*-diol could be obtained during the epoxidation of olefins, utilizing a manganese complex as catalyst and hydrogen peroxide as cooxidant. This process was later enhanced by Boer and colleagues, boosting activity and shifting the selectivity towards *cis*-dihydroxylations.[67] Similar to RuO4, in literature only a few examples for manganese catalyzed AD are reported, mainly owed to common overoxidation and epoxide formation. One example for manganese catalyzed highly enantioselective *cis*-dihydroxylation is the work of Chow and colleagues.[68] Here, they developed a practical protocol for AD of

olefins with potassium peroxymonosulfate (Oxone®), catalyzed by a manganese complex bearing a chiral tetradentate N_4-donor ligand, resulting in improved *cis*-diols generation, reaching up to 95% yield with up to 96% *ee*.

During the past years, the iron-catalyzed *cis*-dihydroxylation came increasingly into focus. This resulted in a convenient strategy, since iron is not only one of the most abundant metals in the earth crust, which makes it affordable, but also because it is relatively nontoxic.[69] Chen and colleagues in 2002 reported the development of the first non-heme iron catalyzed *cis*-dihydroxylation of olefins.[70] Recently, Borrell and colleagues established a green route for the iron catalyzed *cis*-dihydroxylation of olefins, employing aqueous hydrogen peroxide as cooxidant.[71] Advantageously, the process exhibited shorter reaction times and excellent conversions for various substrates. Two drawbacks of this approach are the lack of enantioselectivity and epoxidation, yielding racemic *cis*-diols and epoxide byproducts. Feng and colleagues reported a fascinating iron-based approach, where they accomplished the iron-catalyzed *cis*-dihydroxylation of the aromatic compound naphthalene.[72] The described process involved the use of a biomimetic iron complex and hydrogen peroxide, achieving to mimic the action of the Rieske non-heme iron enzyme naphthalene dioxygenase. This approach currently represents the sole example of direct chemical *cis*-dihydroxylation of aromatics. In contrast to the enzymatic reaction, the generated *cis*-1,2-dihydro-1,2-naphthalenediol is obtained as racemic mixture. To date, literature provides only a sparse number of asymmetric examples of iron catalyzed *cis*-diyhdroxylation. One of the most attractive routes for AD was developed by Zang and colleagues in 2016.[73] Here, a highly enantioselective iron-catalyzed reaction with hydrogen peroxide as oxidant was reported, enabling asymmetric *cis*-dihydroxylation of olefins with a remarkable 99.8% *ee* and a good 85% isolated yield. Therefore, the conversion of 24 different (*E*)- and terminal olefins could be enabled by utilizing various iron complexes bearing tetradentate N_4-ligands (Figure 5).

A

B

Cat = [Fe^{II}(L3)(OTf)$_2$]
Chiral iron(II) catalyst
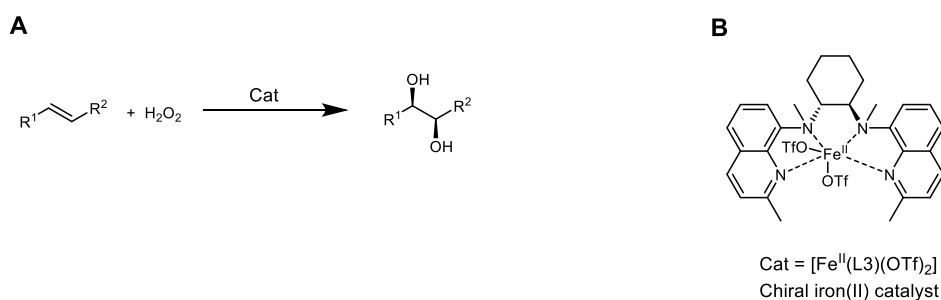
Figure 5: Iron catalyzed *cis*-diyhdroxylation. (**A**) Iron(II) catalyzed asymmetric *cis*-dihydroxylation of an olefin, using hydrogen peroxide (H_2O_2) as oxidant. (**B**) Example for a chiral iron(II) catalyst, containing a tetradentate N_4-ligand. Adapted from Zang and colleagues.[73]

Introduction

Despite the promising results for iron-catalyzed *cis*-dihydroxylations of olefins, their application in organic synthesis has been hampered mainly by the iron(II)-mediated degradation of hydrogen peroxide (Fenton reaction),[74,75] which is frequently used as cooxidant in iron catalysis. Furthermore, the undesired generation of highly reactive hydroxyl radicals can induce decomposition of the substrate, product, and ligand. Nevertheless, despite catalysts based on ruthenium, manganese or iron have shown a better potential for AD, OsO_4 remains as the superior catalyst for *cis*-dihydroxylations.[37]

1.3.3   *Hypervalent iodine mediated* cis-*dihydroxylation*

Hypervalent iodine compounds are attractive alternatives for the metal-free dioxygenation of olefins, owed to their low toxicity and ready availability.[40] In 1958, Woodward and colleagues reported the Woodward reaction, an attractive alternative for the *cis*-dihydroxylation of olefins, employing stoichiometric amounts of silver acetate, iodine and wet acetic acid.[76] Based on this reaction, Emmanuvel and colleagues developed a catalytic and transition-metal-free alternative.[77] Thereby, the dihydroxylation is catalyzed by lithium bromide and mediated by sodium periodate in acetic acid. This reaction works for electron rich, as well as for electron deficient olefins, generating the corresponding *cis*-diols in excellent yield and good diastereoselectivity (Figure 6).
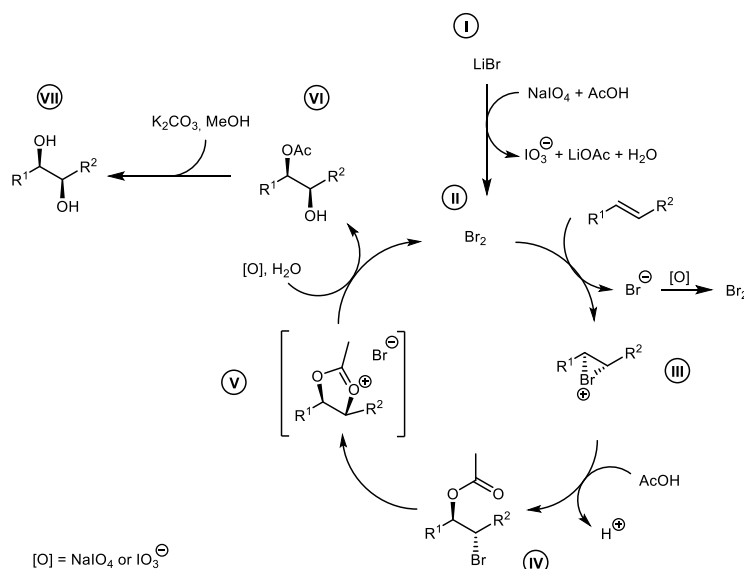
Figure 6: Proposed catalytic cycle for hypervalent iodine Woodward dihydroxylation reaction. Oxidation of lithium bromide (**I**) in addition with $NaIO_4$ generates Br_2 (**II**) *in situ*. Compound (**II**) rapidly undergoes bromoacetoxylation with the olefin *via* a bromonium ion (**III**) to produce the *trans*-1,2-bromoacetate derivative (**IV**). Compound (**IV**) in the presence of $NaIO_4$ forms a 1,3-dioxolan-2-yl cation species (**V**) by an intramolecular cyclisation. Further hydrolysis of (**V**) gives a *cis*-hydroxy acetate entity (**VI**) and Br_2 (**II**). Compound (**VI**) is deacetylated by the addition of K_2CO_3 in MeOH, yielding the corresponding *cis*-diol (**VII**). Adapted and customized from Emmanuvel and coworkers.[77]

11

Zhong and colleagues performed the further enhancement of the Woodward dihydroxylation reaction, enabling reactions at room temperature *via* the use of a (diacetoxyiodo)benzene/boron trifluoride etherate system.[78] Furthermore, the group of Fujita reported an asymmetric variant of the Woodward reactions, employing chiral hypervalent iodine(III) reagents.[79] Thus, enantioenriched (88-96% *ee*) *cis*-diols were generated in good yield. In addition, the enantiomeric excess of this asymmetric reaction could be simply switched by altering the reaction conditions.

It is important to stress that despite all the extraordinary achievements regarding the chemically driven synthesis of *cis*-dihydroxylated compounds, the dearomatizing *cis*-dihydroxylation of aromatics still cannot be performed chemically in one single step,[41] with the exception of the previously mentioned work of Feng and colleagues.[72] Even for the generation of one of the most simple *cis*-diols, *cis*-dihydrocatechol, the current process relies on the high-priced 1,3-cyclohexadiene as starting substrate and involves a three step process, giving a 50% overall yield.[80] Thus, the absence of chemically synthetic approaches that directly access low-cost aromatics as starting materials for the generation of *cis*-dihydroxylated compounds, is yet to be surmounted. Astonishingly, biocatalysis does.

1.4   Biocatalytically driven *cis*-dihydroxylations

Despite the metal-based AD of olefins is well established, its low sustainability in terms of toxicity, environmental impact, and byproduct formation are significant drawbacks. Besides, the enantioselective *cis*-dihydroxylation of *cis*-disubstituted olefins represent still a major challenge (20-80% *ee*).[27] Moreover, the *cis*-dihydroxylation of C=C double bounds of aromatic compounds is not within the reaction scope of AD reactions.

In this sense, enzymes represent an attractive alternative due to their mild and environmental friendly reaction conditions, high chemo-, regio- and enantioselectivity, as well as high catalytic activity.[81] Among the tremendous number of oxidoreductase enzymes found in nature, only Rieske non-heme iron dioxygenases (ROs) are capable of performing the synthetically relevant and demanding *cis*-dihydroxylation reaction, turning them into useful and valuable biocatalysts.[82] Thereby, ROs are not only able to *cis*-dihydroxylate C=C double bonds for a variety of olefins, but also for aromatic compounds with excellent enantioselectivity (Figure 7).[83]
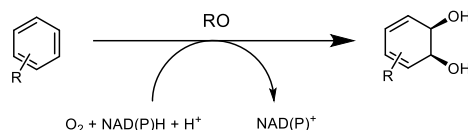
Figure 7: Rieske non-heme iron dioxygenase (RO) -catalyzed dearomatizing *cis*-dihydroxylation of aromatic compounds. The reaction uses molecular oxygen as cosubstrate and NAD(P)H as cofactor.

For instance, toluene dioxygenase (TDO) from *Pseudomonas putida* (*P. putida*) strain F1 is able to perform the dearomatizing *cis*-dihydroxylation of several aromatic substrates, incorporating molecular oxygen directly by harnessing the reductive power of NAD(P)H.[84,85] The following chapter is committed to underline the relevance, scope, and potential of this fascinating biocatalytically-driven reaction with special emphasis on the TDO system.

1.5   Rieske non-heme iron dioxygenases

The conversion of aromatics to nonaromatic molecules is a challenging reaction, due to the strong stability of the aromatic ring-system.[86] Up today, there is only one sole report by Feng and colleagues describing the chemical one step dearomatizing *cis*-dihydroxylation of one single aromatic substrate, yielding a racemic product mixture.[72] Hence, it is highly remarkable how nature evolved enzymes such as ROs, which are capable of catalyzing this highly demanding reaction not only with excellent catalytic activity, but also with outstanding regio- and enantioselectivity (>98% *ee*), for a wide variety of substrates.[1,82] The RO family includes hundreds of dioxygenase-type enzymes from different microorganisms.[87] Nevertheless, the aim of this thesis was centered on enhancing the TDO system from *P. putida* F1. The TDO system was selected based on the following three points; First, from a historical point of view, TDO was the firstly described RO. Second, most of the publications from the past decades used TDO as model enzyme to perform studies regarding structural characterization, mechanism elucidation, and substrate scope of ROs. Third, TDO displays high substrate promiscuity, it is highly manageable to perform preparative biotransformations, and its crystal structure is already available.

1.5.1   Historical background

The fascinating superfamily of ROs was discovered in 1968 by Gibson and colleagues, in a work describing the oxidative degradation of the aromatic compound benzene and its derivatives by the soil bacterium *P. putida* F1.[88,89] In this pioneering work, Gibson showed that *P. putida* F1 is capable of growing on toluene as sole carbon source and to oxidized several

aromatics to their corresponding catechols, specifically *via* the generation of *cis*-dihydrocatechols. Thus, *P. putida* F1 is not only able to tolerate high concentrations of aromatics, toxic to the majority of microorganisms,[90] but also to utilize this compounds as sole carbon source, subsequently metabolizing them into the intermediates of the tricarboxylic acid cycle.[91] Later on, the generation of the mutant strain *P. putida* 39/D enabled the accumulation of *cis*-dihydrocatechols and abolished the secondary reaction to the corresponding catechols completely.[92] In another work, by performing experiments with isotopic $^{18}O_2$, it was revealed that both atoms of one oxygen molecule are incorporated into one substrate molecule.[93] In 1977, Yeh and colleagues achieved to resolve the enzymatic system catalyzing the *cis*-dihydroxylation, into three protein components (reductase, ferredoxin and oxygenase) and named the multicomponent system TDO, due to its participation in toluene degradation.[94] One decade after, in 1988 the group of Gibson elucidated the genetic organization of the chromosomal toluene dioxygenase operon (*todC1C2BA*) from *P. putida* F1, encoding the upper degradation pathway of aromatics.[95] Thereby, the first four steps in the pathway involve the sequential action of toluene dioxygenase TDO (also named TodC1C2BA), *cis*-toluene dihydrodiol dehydrogenase TodD, 3-methylcatechol 2,3-dioxygenase TodE, and 2-hydroxy-6-oxo-2,4-heptadienoate hydrolase TodF. Later, the group of Gibson could elucidate the remaining three genes of the *tod* pathway downstream from *todE*, encoding the 2-hydroxypenta-2,4-dienoate hydratase TodG, the 4-hydroxy-2-oxovalerate aldolase TodH and the acetaldehyde dehydrogenase TodI.[91] In addition, the open reading frame *todX* located upstream of the *todC1* gene, was identified.[96] This gene encodes for the membrane protein TodX, likely involved in substrate transportation. Thereby, a toluene-inducible promoter is localized in front of *todX*, responsible for the expression of all *tod* structural genes (Figure 8A). Thus, the transcription order for the whole *tod* gene cluster is *todXFC1C2BADEGIH* (Figure 8B).[97,98]

A

B
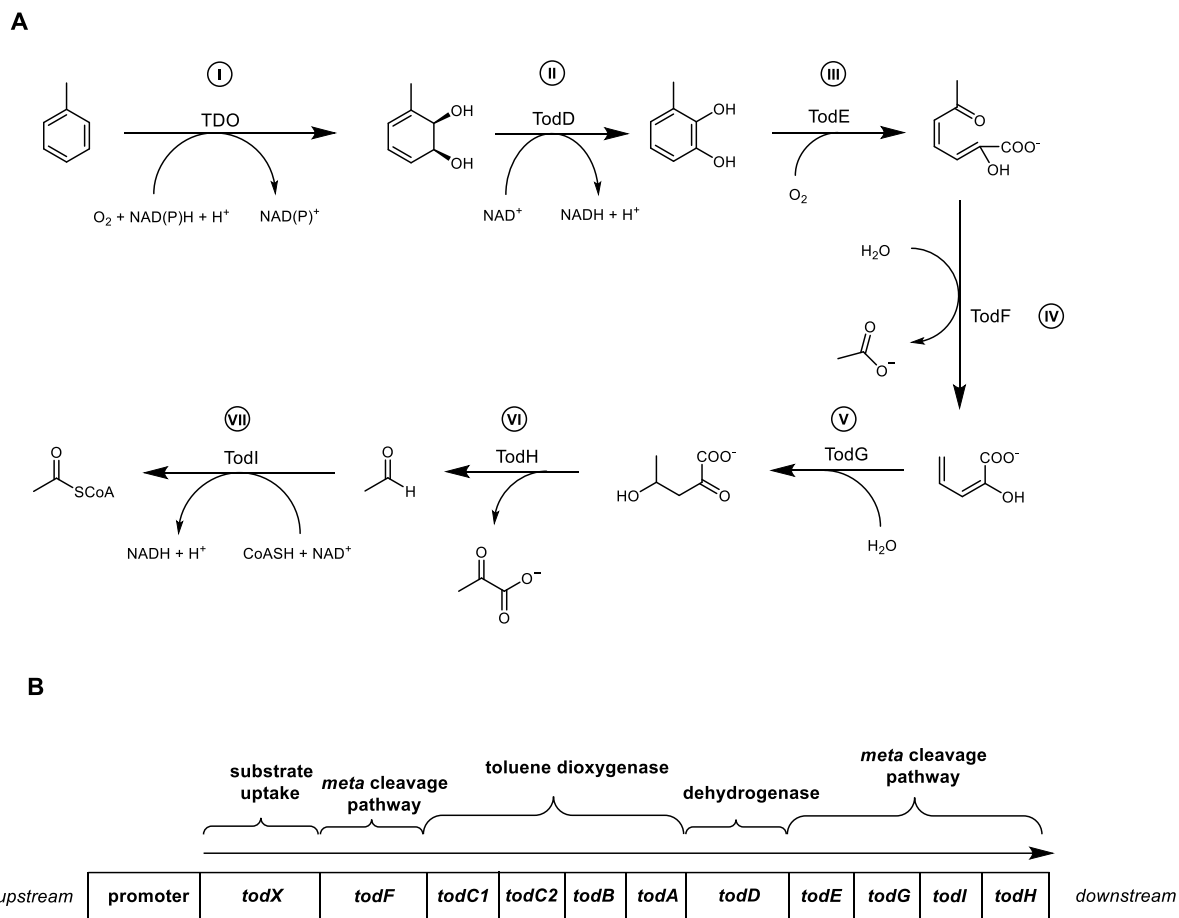
Figure 8: Toluene degradation by *P. putida* F1 and involved *tod* genes. (**A**) Reduction of toluene dioxygenase TDO (**I**), also referred as TodC1C2BA, by NAD(P)H enables the conversion of toluene to *cis*-1,2-dihydro-3-methylcatechol, whereby O_2 is incorporated. Oxidation of *cis*-toluene dihydrodiol dehydrogenase TodD (**II**) by NAD^+ leads to the oxidation of *cis*-1,2-dihydro-3-methylcatechol to 3-methylcatechol, which is further oxidized by the 3-methylcatechol 2,3-dioxygenase TodE (**III**) to 2-hydroxy-6-oxo-2,4-heptadienoate. The later compound is then hydrolyzed by the 2-hydroxy-6-oxo-2,4-heptadienoate hydrolase TodF (**IV**) to 2-hydroxypenta-2,4-dienoate and acetate. Subsequently, hydration takes place *via* the 2-hydroxypenta-2,4-dienoate hydratase TodG (**V**), yielding 4-hydroxy-2-oxovalerate, which is further degraded by the 4-hydroxy-2-oxovalerate aldolase TodH (**VI**) to pyruvate and acetaldehyde. In the final step, acetaldehyde is converted by the acetaldehyde dehydrogenase TodI (**VII**) to acetyl-CoA by employing Coenzyme A (CoASH) and reducing NAD^+. Adapted from Lau and colleagues.[91] (**B**) Illustration of the whole *tod* gene cluster. The promoter is located upstream *todX*. The complete *tod* operon is transcribed in the order *todXFC1C2BADEGIH*. Adapted from Parales and coworkers.[98]

The TDO multicomponent system comprises the flavoprotein reductase TodA, the ferredoxin TodB and the oxygenase. The latter consists out of the α- and the β-subunits TodC1 and TodC2, respectively. Regarding the mechanism, initially reductase TodA accepts two electrons from NADH and transfers them, one by one to the small protein ferredoxin TodB. Then, ferredoxin TodB shuttles the electrons further to the oxygenase TodC1C2. In this way, the reduced oxygenase TodC1C2 catalyzes the *cis*-dihydroxylation of toluene, by incorporating molecular oxygen (O_2), generating the corresponding *cis*-dihydrodiendiol (Figure 9).

Introduction
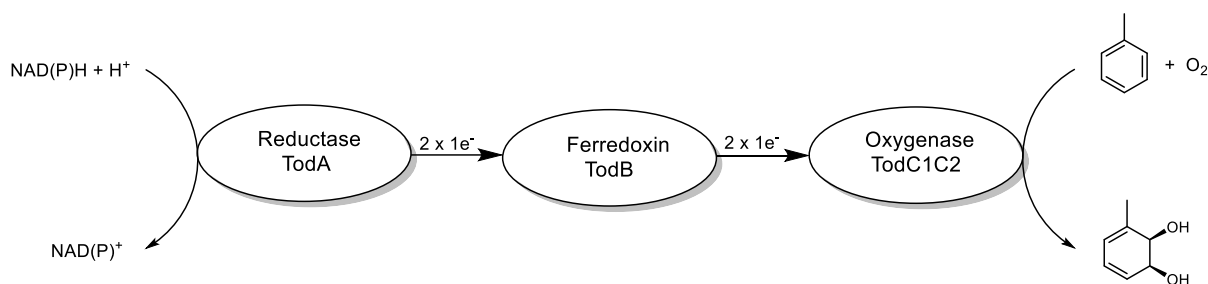
Figure 9: Oxidation of toluene to *cis*-toluene dihydrodiol by the toluene dioxygenase (TDO) multicomponent system. NAD(P)H is reduced by the reductase TodA, which transfers the two electrons one by one to ferredoxin TodB. Then, ferredoxin shuttles the electrons, again one by one, to the oxygenase TodC1C2, where molecular oxygen is incorporated into toluene, to yield *cis*-1,2-dihydro-3-methylcatechol.

One year later, the group of Gibson achieved the heterologous expression of the TDO multicomponent system in *Escherichia coli* (*E. coli*) JM109, by constructing clones harboring the TDO genes *todC1C2BA*.[99] The resulting recombinant organism *E. coli* JM109(pDTG601) was able to overproduce the TDO enzymatic complex, enabling for the first time the *cis*-dihydroxylation of aromatics in the model bacterium *E. coli*. After the TDO characterization by Gibson and coworkers, dating over half a century ago, a multitude of other ROs have been discovered and characterized. Currently, several hundreds of ROs have been described in literature, which impelled the need to establish a classification system.[87] An initial classification of ROs was proposed by Batie and colleagues in 1991, based on the component number and the nature of the redox centers, resulting in the division of three major groups.[100] However the extent of this classification turned out to be limited, since over the last decades many novel ROs have been discovered which could not be classified in any group of the classification suggested by Batie. Therefore, the need to design a new classification system for these enzymes arose. Thus, Kweon and coworkers proposed in 2008 a new scheme for classification, based on new sequence information and interactions between the different RO components, which resulted in categorizing ROs into five different types.[87] The advantage of this system, compared to previous one, is that ROs are analyzed as a whole and that the classification system is capable of responding dynamically to the growing set of ROs. For instance, TDO is classified as type IV RO, representing a three-component system that consist out of an oxygenase, a [2Fe-2S]-type ferredoxin, and a glutathione reductase-type reductase. More recently, based on the work of Kweon and coworkers, Chakraborty and colleagues suggested in 2012 a modified and expanded classification system.[101] In order to understand the fascinating catalytic activity of ROs, further works focused on obtaining the crystal structure aiming to provide deeper insights regarding the system conformation and catalytic mechanism.

1.5.2   *Structure and catalytic mechanism*

A common method for the structural elucidation of enzymes, which also applies for ROs, is the analysis of their crystals *via* X-ray diffraction. In this manner, in 2009 Friemann and colleagues achieved to obtain the crystal structure of TDO (PDB: 3EN1), enabling the identification of active site amino acids and the determination of the oxygenase subunit composition.[102] In general, ROs are two- to three-component systems which consists out of a reductase, an oxygenase, and optionally a ferredoxin protein.[103] The oxygenase is composed of three α-subunits of the form α_3, which in some ROs, such as TDO, are additionally stabilized by three β-subunits to form an $\alpha_3\beta_3$ heterohexamer.[102] For the majority of ROs, the purpose of the oxygenase β-subunit is purely structural. However, some studies have shown that the β-subunit can have influence on the substrate affinity.[104] The oxygenase α-subunit harbors the Rieske cluster domain, eponymous for this enzyme superfamily, and the catalytic domain containing the mononuclear iron (Figure 10).
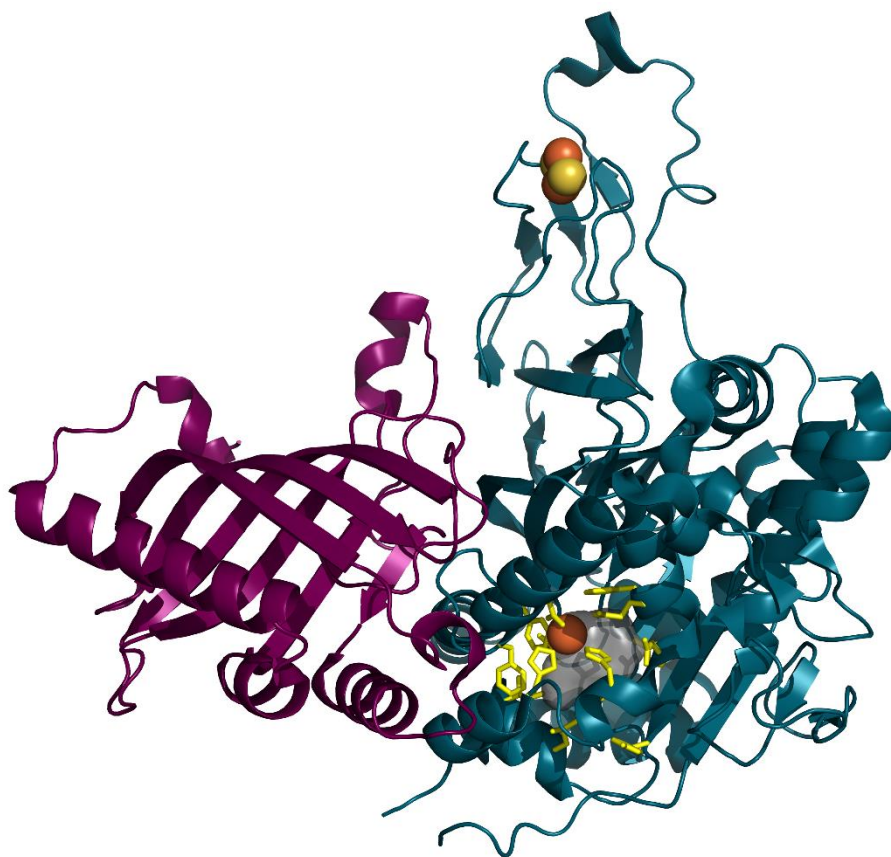
Figure 10: Crystal structure of toluene dioxygenase (TDO) from *P. putida* F1 (PDB: 3EN1).[102] One α- and one β-subunit (dark-cyan and purple, respectively) of the hexameric $\alpha_3\beta_3$ oxygenase enzyme are depicted. The Rieske [2Fe-2S] cluster (S; yellow sphere, Fe; orange sphere) are illustrated in the upper part of the figure. The catalytic iron (orange sphere) is located in the lower figure part, surrounded by 18 active site residues (yellow sticks), conforming the active site (dark grey spherical shape). The structural model was generated with the software PyMOL, employing the TDO crystal structure.

Introduction

The Rieske center is a rhombohedral [2Fe-2S] iron-sulfur cluster, bound to the protein *via* two cysteine-sulfur atoms and two histidine-nitrogen atoms (Figure 11A).[105] The cysteine-bound iron(III) retains its trivalence at all times, whereas the histidine-bound iron(III) can be reduced to iron(II). The Rieske cluster is commonly found in proteins involved in electron transport or the respiratory chain and serve for electron transfer.[106] In ROs, the Rieske cluster transfer the electrons received from the ferredoxin further to the catalytic mononuclear iron in the active site. The distance between the Rieske cluster and the catalytic iron to reach a single α-subunit is to long for viable electron transfer (~45 Å).[103,107] Nevertheless, the distance is considerably shorter (~12 Å) between the Rieske cluster of one α-subunit and the catalytic iron of an adjacent α-subunit in the α₃ or α₃β₃ quaternary structure configuration. The α-subunits are directly connected by hydrogen bonds through a highly conserved aspartic acid residue, enabling the electron transfer (Figure 11B). The catalytic mononuclear iron is coordinated in many ROs by one water molecule, and in a facial triad motif of two histidines and one carboxylic acid (aspartate or glutamate), which is a common way to incorporate an iron(II) at the catalytic site of these proteins.[103] This coordination type leaves the face of the iron exposed to the large, hydrophobic active site of ROs, creating a catalytic platform where molecular oxygen can bind and react with the substrate.
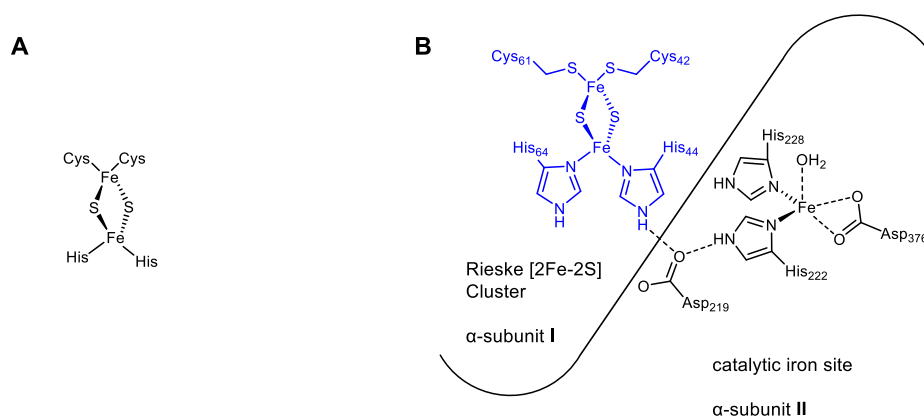
Figure 11: (**A**) Rieske-type [2Fe-2S] iron-sulfur cluster. One iron (Fe) is coordinated by two cysteines, the other by two histidines. Both iron atoms are coordinated by two sulfur atoms (S). (**B**) Route of electron transfer between adjacent α-subunits in toluene dioxygenase (TDO). The Rieske center (blue) in α-subunit **I** is connected to the catalytic center of the adjacent α-subunit **II**. The connection is through the residue Asp₂₁₉, which is hydrogen bonded to a His₄₄ residue of the Rieske center and His₂₂₂ residue at the active site. Adapted from Kauppi and coworkers and customized for toluene dioxygenase.[107]

Till today, the knowledge of the RO mechanism is relatively limited due to the difficulty in obtaining spectroscopic information on the mononuclear catalytic iron in the presence of the oxygen sensitive Rieske cluster.[108] Moreover, the catalytic iron is lacking a significant chromophore and contains an electron paramagnetic resonance-silent ferrous (d^6) iron atom in the resting state, hindering spectroscopic studies.[109] During the last two decades, various

publications aimed for the elucidation of the catalytic RO mechanism, by spectroscopic approaches, as well as by quantum mechanical modelling.[110–112] One of the ROs with the best studied mechanism for the *cis*-dihydroxylation, is naphthalene dioxygenase (NDO) from *Pseudomonas species* NCIB 9816-4. The crystal structure of NDO was solved in 1998 (PDB: 1O7N),[107] revealing the side-on binding of the molecular oxygen to the catalytic iron.[113] Furthermore, single turnover chemistry and near-infrared magnetic circular dichroism spectroscopy were just two strategies utilized for the revelation of NDO mechanism.[114–116] Considering the homology between NDO and several other ROs, for instance TDO (47% amino acid sequence similarity to NDO), the described mechanisms can be consider equivalent for such systems (Figure 12). Despite the numerous literature reports, there is still not complete unanimity regarding the mechanism, which derives in the existence of two slightly different ones.

The first one involves a Fe(III)-OOH complex as reactive species, the second one a Fe(V)=O(OH) complex.[109,114] Both mechanisms are initiated by the binding of the substrate to the active site, altering the coordination geometry of the catalytic iron complex, by the elimination of a water molecule from a distorted octahedral to a square pyramidal geometry.[114] Subsequent reduction of the Rieske cluster affords a mixture of square pyramidal and trigonal bipyramidal coordination geometries for the catalytic iron.[108] These modifications in geometry are originated from the changes in protein conformation, altering the location of the coordinating ligands and the catalytic iron. Thus, space in the active pocket is created, enabling molecular oxygen to bind to the catalytic iron. This means that molecular oxygen exclusively binds after the substrate is bound and the Rieske center has been reduced, which prevents oxygen activation in the absence of substrate and therefore the formation of reactive oxygen species.[109] The side-on binding of molecular oxygen is a unique feature of ROs, which allows each oxygen atom to attack the C=C double bound from the same face of the substrate, resulting at the end into a *cis*-dihydroxylated product.[113] The side-on binding of the molecular oxygen is followed by the transfer of one electron from the reduced Rieske center to the catalytic iron, yielding a ferric peroxide complex. By subsequent protonation and the loss of a water ligand, the complex is transformed into a bidentate hydroperoxide complex Fe(III)-OOH.[117] As previously mentioned, for the next reaction step two different mechanisms are proposed in literature. The first one involves a coupled O-O bond cleavage and substrate oxidation, yielding a Fe(IV)=O complex and a hydroxylated substrate radical intermediate, which reacts further to a ferric alkoxyhydroxy substrate complex. In the second proposed mechanism, O−O bond cleavage takes place prior to substrate oxidation, producing a Fe(V)=O(OH) complex, which

reacts with the substrate to form the ferric alkoxyhydroxy substrate complex. The transmission of a second electron from the Rieske cluster to the catalytic iron, results in the formation of a ferrous alkoxyhydroxy substrate complex, which after protonation yields the regenerated resting state of the oxygenase and the *cis*-dihydroxylated product.[109]

It is worth to mention, that recent studies of the Lipscomb group proposed a different reaction mechanism for the benzoate 1,2-dioxygenase from *P. putida* mt-2, involving an end-on binding of molecular oxygen and the formation of a Fe(III)−superoxo species.[112,118]
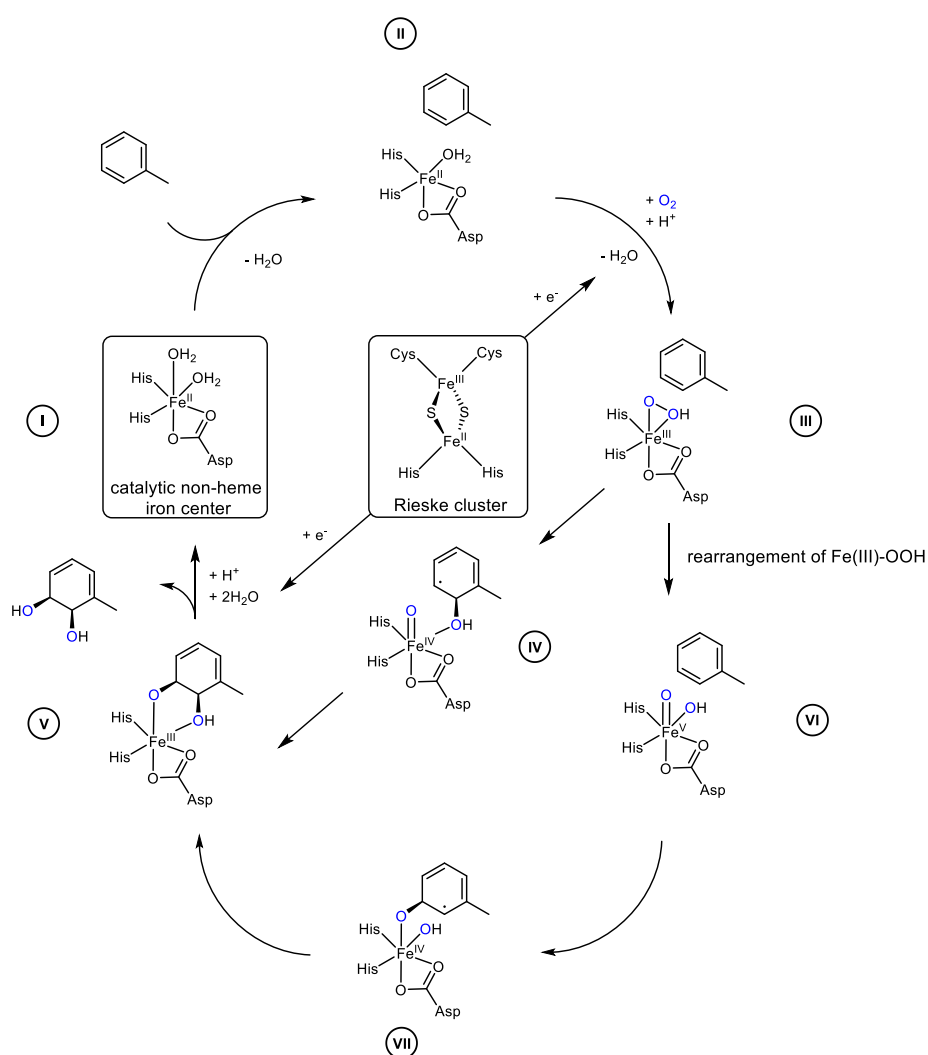
Figure 12: Proposed catalytic mechanism for toluene dioxygenase (TDO) catalyzed *cis*-dihydroxylation of toluene. Binding of toluene to the resting state of the oxygenase (**I**) leads to the loss of one water ligand and to the alteration of the coordination geometry (**II**). This change in geometry enables molecular oxygen (blue) to bind side-on to the catalytic iron. Then, one electron from the reduced Rieske center is transferred to the catalytic iron, where the subsequent protonation and water ligand elimination yields the bidentate hydroperoxide complex Fe(III)-OOH (**III**). A coupled O-O bond cleavage along with substrate oxidation yields complex Fe(IV)=O and the hydroxytoluene radical intermediate (**IV**), which reacts further to the ferric alkoxyhydroxytoluene complex (**V**). Alternatively, O-O bond cleavage of (**III**) takes place prior to substrate oxidation, producing the Fe(V)=O(OH) complex (**VI**), which reacts with toluene to form the hydroxytoluene radical intermediate (**VII**) and subsequently the ferric alkoxyhydroxytoluene complex (**V**). The transmission of a second electron from the Rieske cluster to the catalytic iron results in the formation of the ferrous alkoxyhydroxytoluene complex, which after protonation yields the regenerated resting state of the oxygenase (**I**) and the product *cis*-(1*S*,2*R*)-1,2-dihydro-3-methylcatechol. Adapted from Barry and Challis and customized for the substrate toluene.[109]

Introduction
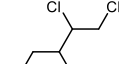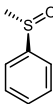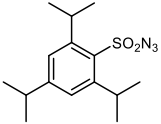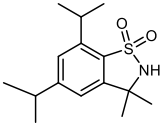
The knowledge behind the catalytic mechanism and the availability of crystal structures, opened the door to have a look inside the outstanding flexibility of the active pocket and therefore to the fascinating large substrate scope, and even more, to the striking reaction spectrum of TDO.

1.5.3   Substrate and reaction spectrum

ROs are key enzymes involved in bacterial catabolism of aromatic compounds, catalyzing the initial dearomatizing *cis*-dihydroxylation of aromatic rings.[103,119] Thus, the majority of ROs exhibit a very broad substrate scope, showing high promiscuity towards non-natural substrates ranging from monocyclic aromatics, heterocyclic and polycyclic arenes to olefins.[120] For instance, the extensively studied TDO is able to hydroxylate over 100 different non-natural substrates.[1,84] In addition to their broad substrate scope, ROs are capable of performing a striking variety of different reactions. For instance, TDO is not only able to perform *cis*-dihydroxylations, but also monohydroxylations,[121] *O*-dealkylations,[122] desaturations,[123] sulfoxidations,[124] and C-H aminations[125] (Table 2). Hereinafter, some selected TDO-catalyzed reactions are given; as previously mentioned, the native reaction of TDO is the *cis*-dihydroxylation of monocyclic aromatic compounds. For this particular reaction numerous reports have been published. For instance, the described conversion of toluene **11** to *cis*-(1*S*,2*R*)-1,2-dihydro-3-methylcatechol **12** with high product formation (95%) and outstanding enantioselectivity (>98%).[84,92,126] Another notable report was made by Lange and colleagues, elaborated on the *cis*-dihydroxylation of multiple aliphatic olefins, like 3,4-dichloro-1-butene **13** which was converted to 3,4-dichlorobutane-1,2-diol **14**.[127] Furthermore, Boyd and colleagues investigated the TDO-catalyzed α-allylic monohydroxylation, employing benzocycloalkenes, e.g. indane **15**, which yielded enantiopure (*R*)-1-indanol **16**.[121] It has been also shown that *O*-dealkylations can be catalyzed by TDO, this enabled the biotransformation of anisole **17**, generating minor amounts of phenol **18**.[122] Another quiet interesting reaction is the dehydrogenation, which occurs for instance in the conversion of 1,2-dihydronaphthalene **19**, yielding naphthalene **20**.[123] Boyd and colleagues investigated the TDO-catalyzed sulfoxidation, by using the substrate thioanisole **21** to generate enantiopure (*R*)-methylsulfoxide **22**, not only in high yields, but also enantiomerically pure.[124] Recently, Vila and colleagues described the C-H amination of the compound 2,4,6-triisopropylbenzenesulfonyl azide **23**, producing 5,7-diisopropyl-3,3-dimethyl-2,3-dihydrobenzo[*d*]isothiazole 1,1-dioxide **24**.[125]

Table 2: Selected examples of the different reaction types catalyzed by toluene dioxygenase (TDO) from *P. putida* F1.

Reaction type	Substrate	Product	Product formation [%] / *ee* [%]	Literature reference
cis-dihydroxylation of aromatics	**11**	**12**	95 / >98	[126]
cis-dihydroxylation of olefins	**13**	**14**	90 / nd	[127]
monohydroxylation	**15**	**16**	66 / >98	[121]
O-dealkylation	**17**	**18**	1 / -	[122]
dehydrogenation	**19**	**20**	5 / -	[123]
sulfoxidation	**21**	**22**	90 / >98	[124]
C-H amination	**23**	**24**	7 / -	[125]

1.5.4   Applications

In general, RO-catalyzed biotransformations generate enantiomerically pure *cis*-dihydro-diendiols when employing aromatic compounds as substrates, yielding synthons of high synthetically value.[24,82] Thereby, these *cis*-diols possess some unique features, such as a proenantiotopic symmetry plane (Figure 13), allowing their enantiodivergent functionalization to yield both enantiomers of one product.[128] In addition to functionalization or substitution of the alcohol groups, the two chiral centers can be utilized to direct stereo-controlled Michael addition reactions (conjugate addition) to, or Diels-Alder reactions (cycloaddition) at, the C=C double bonds of the cyclohexadiene moiety.[129] Furthermore, the C=C double bounds can be employed further for selective epoxidation, aziridation, and ozonolysis.[82]
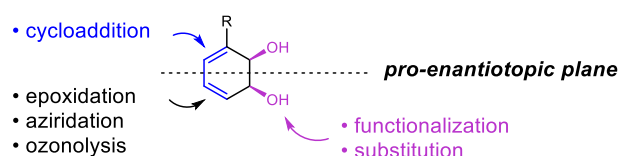
Figure 13: Possible functionalization of *cis*-dihydrodiendiols. The pro-enantiotopic plane is introduced as dotted line. Adapted from Özgen and Schmidt.[82]

Consequently, enantiopure *cis*-dihydroxylated aromatics are commonly used for the preparation of natural products and pharmaceutical compounds.[1] The unique synthetic potential of ROs is of great interest and demand for actual application, since there are no chemically based synthetic approaches enabling the enantioselective *cis*-dihydroxylation of aromatics. Therefore, ROs are receiving an increasing interest as biocatalysts in organic synthesis owed to their broad substrate scope, high yield, and excellent enantioselectivity.[82,130] Especially, the pioneering work of the Hudlicky group strongly emphasizes the application of TDO-catalyzed *cis*-diols as chiral synthons for the synthesis of natural products.[24,84] The combination of ROs with traditional synthetic chemistry led to unprecedented levels of efficiency for the total synthesis of some complex molecules.[131–133] To circumvent the multicomponent nature of the TDO system, the external addition of reduction equivalent NAD(P)H,[134] and the observed instability of purified ROs,[94,135] the majority of TDO-based biotransformations are performed in whole cell platforms for better efficiency. These whole cell systems are either native or recombinant unicellular organisms, harboring the RO of interest. The development of such cell factories has enabled the synthesis of sought-after enantiomerically pure *cis*-dihydrodiendiols (Table 3). For instance, TDO-catalyzed whole cell biotransformation of benzene **25** yields the simple, though valuable synthon *cis*-1,2-dihydro-catechol **26**,[93] which can be utilized for the total synthesis of the pharmaceutically relevant
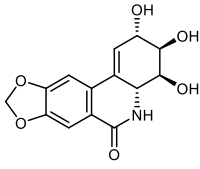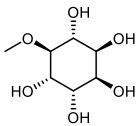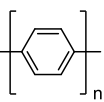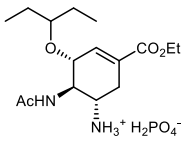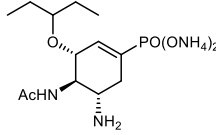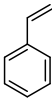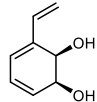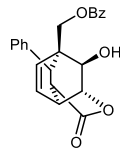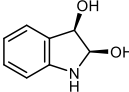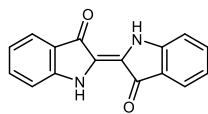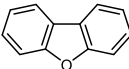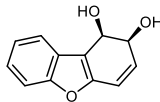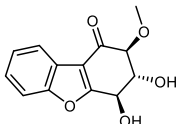
molecules (+)-lycoricidine **27**,[131] and (±)-pinitol **28**,[136] as well as for the polymer polyphenylene **29**.[137] The broad scope of ROs applications is so tangible up to the point that these biocatalyst have even been subject of patents, like in the case of the patented RO-catalyzed biotechnological route to various *cis*-dihydroxylated monocyclic aromatics by Imperial Chemical Industries.[138] For instance, the TDO-driven biocatalytic conversion of bromo-benzene **30**, generates the synthon *cis*-(1*S*,2*S*)-1,2-dihydroxy-3-bromocyclohexa-3,5-diene **31**,[139] which can be further utilized for the synthesis of the two anti-influenza agents Tamiflu® **32** and tamiphosphor **33**.[140] Moreover, by employing TDO as catalyst, Hudlicky and colleagues established an approach to convert styrene **34** to (1*S*,2*R*)-1,2-dihydroxy-3-vinylcyclohexa-3,5-diene **35**, which was further used for the generation of the bioactive molecule (-)-zeylena **36**.[141] A highly interesting application of ROs is the dihydroxylation of indole **37** to (2*R*,3*R*)-2,3-dihydro-1H-indole-2,3-diol **38**, which spontaneously dehydrates and subsequently dimerizes, yielding the widely utilized blue dye indigo **39**.[142] The profit behind the biocatalytic green route to the prized dye, resulted in a patent claimed by Genencor International in 1995 for the NDO-catalyzed microbial production of indigo **39**.[143] It is worth to mention that this valuable reaction can not only be efficiently catalyzed by NDO, but also by TDO.[144] A recent approach was developed by Boyd and colleagues, enabling the TDO-driven oxyfunctionalization of dibenzo[*b,d*]furan **9** to generate (1*R*,2*S*)-1,2-dihydro-dibenzo[*b,d*]furan-1,2-diol **40**. This accomplishment shortened the synthesis of the fungal metabolite (-)-ribisin-A **10** marginally, though without the need for C-C bond-forming reactions.[36]

The mentioned examples illustrate the vast potential of TDO and their relevant and factual application. Strikingky, ROs are the only biocatalyst on hand to generate directly enantiomerically pure *cis*-dihydrodiendiols from aromatic compounds. Nevertheless, there are still challenges to overcome regarding substrate scope diversification, performance, efficiency, and stability. In this sense, protein engineering approaches are key to develop improved RO variants customized for specific targets.

Table 3: Selected examples to illustrate TDO application in the synthesis of *cis*-dihydroxylated synthons. Thereby, the utilized aromatic substrate, the biosynthesized *cis*-diol synthon, and its further use for the generation of relevant molecules (natural products, pharmaceuticals, fine chemicals) are shown.

Substrate	*cis*- Dihydroxylated synthon	Natural product	Literature reference
		27	[131]
25	**26**	**28**	[136]
		29	[137]
30	**31**	**32**	[140]
		33	[140]
34	**35**	**36**	[141]
37	**38**	**39**	[144]
9	**40**	**10**	[36]

1.5.5 Protein engineering

Protein engineering is inspired in the fundamental principle of evolution with the aim of introducing mutations to boost catalytic activity, to improve enzyme stability, or even to customize protein specificity towards non-natural substrates, pursuing to satisfy the industrial requirements.[145] However, despite their overall high catalytic activity and selectivity, natural ROs are often not suitable for industrial applications. Nevertheless, thanks to the advances in enzyme engineering, nowadays it is feasible to apply this powerful tool in order to enhance certain enzyme features. The established approaches in the field of enzyme engineering have reached an incredible proficiency level. This progress has enabled the alteration of the natural properties of numerous targeted proteins aiming to perform desired reactions at advantageous yields and selectivities.[10,12,14,18] During the past two decades, rational, semi-rational and evolutionary methods have been employed to enhance several sought-after features in ROs. As starting point, the main aspects considered in such mutagenesis strategies were directed on correlate the protein structure and function of the selected RO with the aim of the study. In this sense, most of publications have focused on engineering the catalytic oxygenase α-subunit. For instance, one of the first reported RO mutagenesis studies, was performed in 1996 by Jiang and colleagues.[146] Here, two of the three mononuclear iron ligands, H222 and H228 could be identified, by applying site-directed mutagenesis on highly conserved amino acids of the TDO α-subunit. Later on, in 2009 Friemann and colleagues confirmed these findings, by achieving the crystallization of the heterohexamer TDO oxygenase (PDB: 3EN1), which in addition revealed the third ligand to be D376.[102] In 1998, Beil and coworkers generated, *via* site-directed mutagenesis, hybrid dioxygenases of the homologues TDO and chlorobenzene dioxygenase (TEC) from *Burkholderia sp.* strain PS12.[147] Thereby, TDO variant M220A was identified to be the one performing the dechlorination of tetrachlorobenzene, a reaction that originally was not performed by TDO wild type. Parales and colleagues unveiled the importance of NDO position F352 (corresponding to F366 in TDO), to alter regio- and enantioselectivity significantly, as derived variants at this position showed for the substrates naphthalene, biphenyl, and phenanthrene.[148] One renowned use of directed evolution techniques on TDO, is represented by the work of Zhang and coworkers.[149] Here, TDO was evolved *via* epPCR towards the conversion of indene to *cis*-(1*S*,2*R*)-indandiol, a potential key intermediate in the chemical synthesis of the anti-viral pharmaceutical indinavir **2** (Crixivan®). In this study, the identified outperforming multiple variant displayed not only a significant increase in product formation, but also a high selectivity for *cis*-(1*S*,2*R*)-indandiol. One notable example encompassing directed evolution techniques was performed by the Arnold group,

where they accomplished to increase the TDO-catalyzed conversion of the poorly accepted heterocyclic substrate 4-picoline.[150] This work involved the generation of over 9000 TDO variants and the establishment of a high-throughput screening for activity evaluation, yielding a multiple variant with 5.6-fold higher activity towards the substrate of interest. Another example was provided by Newman and coworkers, which generated *via* multigene DNA shuffling of TDO and TEC a total of 1044 variants and screened the obtained library towards the conversion of *p*-xylene.[151] Their aim was to increase product formation towards the sole product *cis*-1,2-dihydroxy-3,6-dimethyl-3,5-cyclohexadiene, which can be utilized for the synthesis of 4-hydroxy-2,5-dimethyl-2,3-dihydrofuran-3-one, the flavor compound strawberry furanone. The best performing chimeric dioxygenase exhibited, in comparison to the wild type, a 4.4-fold higher activity. More recently, in 2015 the Hauer group reported the site-directed mutagenesis of cumene dioxygenase (CDO) from *P. fluorescens* IP01 towards the conversion of olefins.[152] Thereby, the active site variant M232A (corresponding to M220 in TDO) showed to increase the product formation for multiple substrates, enabling in addition, the conversion of (+)-α-pinene, a substrate not converted by CDO wild type. A notable example regarding TDO rational engineering, was reported by Vila and colleagues, who used a smartly designed set of only three TDO variants at active site positions I324, T365 and F366. [153] By addressing such positions, the chemo-, regio-, and enantioselectivity for various substrates could be influenced substantially. In terms of semi-rational engineering, the Hauer group generated a selected set of variants to investigate their influence at the active site of NDO. The results highlighted six variants, located at the active site positions F202, A206, V260, H295, F352, and L307 (corresponding to F216, M220, L272, V309, F366 and L321, in TDO respectively), to have a tremendous influence in chemo-, regio-, and enantioselectivity for various monocyclic aromatics.[154] Furthermore, just recently, the Hauer group reported that the activity and selectivity of the CDO system could be also influenced by introducing loop variations.[155] Thereby, investigations on libraries generated by alanine scan, insertion, and deletions of loop structures located near the active site, resulted in up to 16-fold higher activities for the substrates styrene, (*R*)-limonene, and 2-phenylpyridine.

Hence, by harnessing the power of protein engineering it is now possible to reshape biocatalyst such as TDO in order to improve catalytic activity and protein stability, to alter chemo-, regio-, and enantioselectivity, and even to confer a completely new substrate scope.

2 MOTIVATION

Rieske non-heme iron dioxygenases (ROs) are fascinating enzymes catalyzing the *cis*-dihydroxylation of aromatic compounds, yielding the corresponding *cis*-dihydrodiendiols. These valuable compounds are employed as synthons in the synthesis of various natural products, pharmaceuticals, and fine chemicals. The RO driven generation of *cis*-dihydrodiendiols is efficient, environmentally friendly, and sustainable. Therefore, the production of *cis*-diols utilizing ROs is superior to the cumbersome chemical-based multi-stage synthesis, which frequently involves toxic metal catalyst and organic solvents. Currently, due to the global pollution and scarcity of resources, the transition to more environmentally friendly and economic synthesis routes is compulsory. Having such premises in mind, the BMBF sponsored project *PowerCart* (031B0369A) was launched. In line with the scope of this project, the objectives of the present thesis aimed to establish and enhance the RO catalyzed *cis*-dihydroxylation of attractive molecules. Thereby, the project was divided into three main stages:

1) Construction of an efficient recombinant RO platform in the model organism *E. coli* and its characterization with regard to the conversion of selected mono- and bicyclic (hetero)aromatic compounds. This starting point was based on certain sought-after features in ROs such as broad substrate spectrum and excellent catalytic activity. For this reason, toluene dioxygenase (TDO) from *P. putida* F1 was selected, since it is one of the best characterized and most widely studied ROs in literature.

2) Generation of a TDO mutant library directed to improve product formation, as well as chemo-, regio-, and enantioselectivity towards bulkier bicyclic (hetero)aromatics. Particularly, for this task, the TDO active site was addressed for the identification of hot-spot positions, substantially influencing the activity and selectivity for different model substrates.

3) Semi-preparative biosynthesis, isolation and characterization of the products, generated by identified outperforming TDO variants. To illustrate the extent of the application of the established and improved platform, the last task was intended to demonstrate the power of enzyme engineering to enable the factual production of substantial amounts of targeted valuable *cis*-diols from non-natural substrates.

3 SUMMARY OF RESEARCH AND DISCUSSION

The derived results of the present dissertation were published in three major research articles (I, II, and III), which are enclosed in chapter 6.[156–158] The first part is focused on the construction and optimization of a suitable, recombinant whole cell RO platform. The developed system resulted to be the heterologous host *E. coli* BW25113 harboring the plasmid TDO-pBAD18. To validate biocatalytic functionality, the platform was challenged for the conversion of the simple substrate benzene, enabling the semi-preparative production of the versatile synthon *cis*-dihydrocatechol in substantial amounts. These results are published in the research article I, identified with the literature reference number [156].[156] The second part is centered on the development of an enhanced TDO platform, towards the conversion of the non-natural substrate naphthalene. To alter enantioselectivity, the TDO active site position F366 was addressed for site-directed mutagenesis. The findings showed that the selective production of both *cis*-1,2-dihydro-1,2-naphthalenediol enantiomers is feasible by applying as biocatalysts TDO wild type and the single point variant TDO_{F366V}. These results are published in research article II identified with the literature reference number [157].[157] The third part, is committed to the generation of a semi-rational designed TDO mutant library, consisting out of 176 single and double variants, aiming to address the active site and the substrate entrance channel of TDO towards the conversion of the bicyclic substrates naphthalene, 1,2,3,4-tetrahydroquinoline, and 2-phenylpyridine. The generated variants were examined regarding their ability to confer changes in chemo-, regio-, and enantioselectivity, as well as in product formation for bulky bicyclic substrates. Successfully, the findings enabled the identification of hot-spot positions positively influencing the sought-after characteristics. Therefore, the most promising variants were utilized for the semi-preparative preparation of the pursued valuable products. These results are published in research article III identified with the literature reference number [158].[158]

The present summary illustrates the extent of the doctoral work, starting with the establishment of a novel and efficient TDO-based platform, followed by the generation of TDO-based biocatalysts, which were redesigned by leveraging the power of enzyme engineering. These achievements allowed to improve the catalytic activity of TDO and also, to opportunely alter its chemo-, regio- and enantioselectivity, resulting in the tangible production of the targeted synthons. In the following sections, the key findings of all these three publications are discussed and aligned with the established objectives described in the motivation chapter.

3.1 Insights in research article I: An enhanced toluene dioxygenase platform for the production of *cis*-1,2-dihydrocatechol in *Escherichia coli* BW25113 lacking glycerol dehydrogenase activity

Presently, a multitude of ROs catalyzing the *cis*-dihydroxylation of (hetero)aromatics have been described in literature.[82] To this extent, one of the best studied RO is TDO from *P. putida* F1, which not only exhibits an impressive substrate scope of over 100 different molecules, but also a high catalytic activity, enabling the preparative synthesis of various interesting products on a gram-scale.[1,84] In addition, the crystal structure of TDO is solved, facilitating the identification of the active site residues and expanding molecular docking studies.[102,159–161] Moreover, previous preliminary studies indicated that the biocatalytic potential of TDO, was suitable to successfully accomplish the objectives of the doctoral investigations addressed in this work. Therefore, in the present dissertation, this remarkable enzyme was selected over other ROs for the generation of valuable *cis*-dihydrodiols. Regarding the heterologous host, *E. coli* was chosen for protein production, since it is one of the best established cell factories, displaying fast growing rates, reaching high theoretical cell density limits, and favoring the controlled protein overexpression. [162]

Indeed, in 1989 Zylstra and colleagues generated the recombinant *E. coli* system JM109 (pDTG601), harboring the toluene dioxygenase genes *todC1C2BA*.[99] Since its development, JM109 (pDTG601) was frequently employed for the generation of *cis*-dihydroxylated synthons with good yield.[84,139,153,163,164] Nevertheless, this TDO platform exhibits multiple drawbacks, for instance the high sensitivity of *E. coli* JM109 to high glucose concentrations as well as its high production level of acetate.[165] In addition, the plasmid expression of pDTG601 is under the *tac* promotor, which activation is induced by the rather expensive inductor isopropyl β-D-1-thiogalactopyranoside (IPTG). Attempts were made to enhance the plasmid system, resulting in the vectors pKST11 and pSPM01, both harboring the genes *todC1C2BA*.[166,167] Nevertheless, both plasmid systems were still designed under the *tac* promotor and heterologously expressed in *E. coli* JM109, serving as cell factory. Thus, in the scope of the project *PowerCart*, a more suitable and attractive TDO platform for industrial purposes was developed. This task was performed by the postdoctoral researcher Dr. Julia Halder.[156] The new design was based on the tightly regulated high-level expression plasmid pBAD18,[168] which offers the advantage of being induced by the much cheaper sugar arabinose, compared to IPTG. Thus, the construction was performed *via* Gibson assembly[169], using the pBAD18 backbone and the four *todC1C2BA* genes as inserts (Figure 14). Thereby the order of the *todC1C2BA* genes was maintained as in the natural host *P. putida* F1. For appropriate heterologous expression, *E. coli* BW25113 was selected as host strain, since this genetic

Summary of Research and Discussion

background lacks the *araBAD* operon and thus cannot metabolize the inductor arabinose.[170] An additional benefit of selecting BW25113 as cell factory is that this strain was utilized to establish the KEIO collection.[171] The KEIO collection comprises individual single-gene deletions of all non-essential genes of *E. coli* BW25113, offering a helpful resource for the systematic analysis of unknown gene functions.
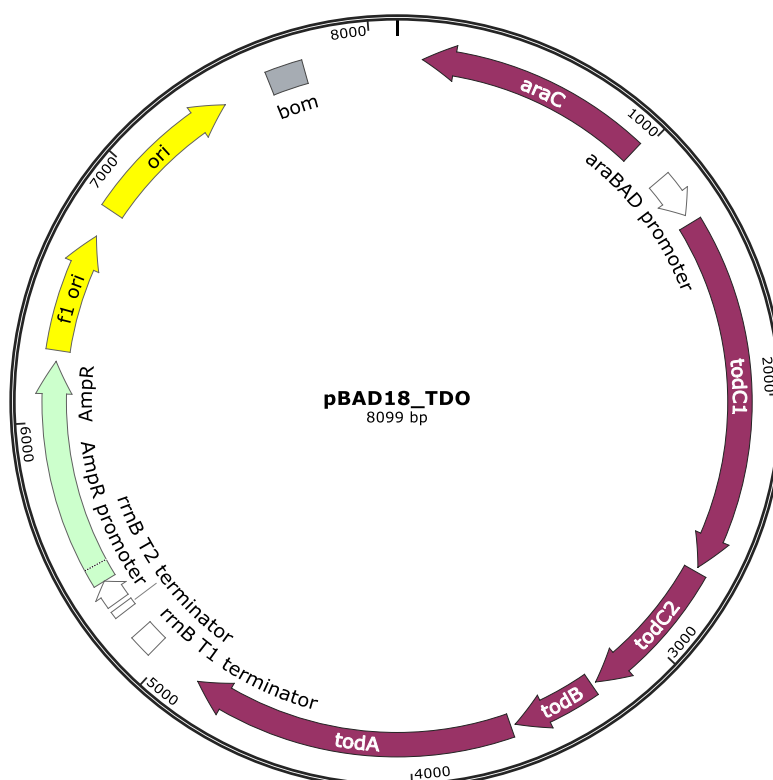
Figure 14: Plasmid of pBAD18-TDO. The plasmid pBAD18 harbors the toluene dioxygenase genes in the same order as in the natural host; oxygenase α-subunit (*todC1*), oxygenase β-subunit (*todC2*), ferredoxin (*todB*), and reductase (t*odA*) under the control of an arabinose promotor (*araC*). The resistance marker of the plasmid is encoded by an ampicillin cassette (*AmpR*). The plasmid card was generated with the SnapGene© GraphPad software.

To evaluate the activity of the newly generated TDO platform, the simple monocyclic aromatic molecule benzene **25** was selected as model substrate.[156] Substrate **25** has been previously reported in multiple publications to be converted by TDO into *cis*-dihydrocatechol **26** as sole product (Figure 15).[86,93,138,166,172]
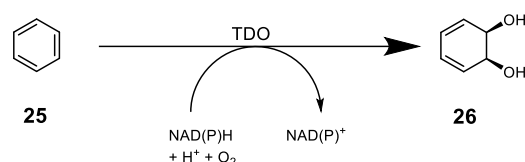
Figure 15: TDO-catalyzed conversion of benzene **25** to *cis*-dihydrocatechol **26**, using molecular oxygen as cosubstrate and NAD(P)H as cofactor. Adapted from Wissner and colleagues.[156]

Despite its simplicity, *cis*-Diol **26** is a versatile synthon since it can be utilized for the synthesis of pharmaceutical compounds of interest such as, (+)-lycoricidine **27**[131] and (±)-pinitol **28**,[136] as well as for the production of the polymer polyphenylene **29**,[137] and precursors of organic semiconductors.[173] Chemically, this valuable synthon can be accessed in a three step process, using 1,3-cyclohexadiene as starting substrate with an overall yield of 50%.[80] Biocatalytically, the native TDO host organism *P. putida* F1, as well as derived mutants, and recombinant organisms expressing TDO, have been used to generate compound **26** in good yield as shown in literature.[84,86,166] However, in the current study, when the *E. coli* BW25113 resting cells harboring pBAD18-TDO were tested, surprisingly catechol **41** and not the expected compound **26** was obtained as main product.[156] To shed light on the so far unreported formation of **41**, Julian Ludwig characterized the influence of various protein expression conditions and working reaction parameters in a complimentary study that resulted in a master thesis.[174] The results showed that by altering the protein expression temperature and some operational reaction parameters such as, reaction time, pH, substrate concentration, substrate feeding, and reaction temperature, the product ratio of **26** to **41** could be beneficially influenced. Especially, protein expression temperature and reaction time exhibited the major influence in product distribution. As observed, with longer reaction times the concentration of the undesired product **41** increased, whereas the concentration of **26** decreased, indicating a degradation of the latter into the aromatic compound **41** (Figure 16).
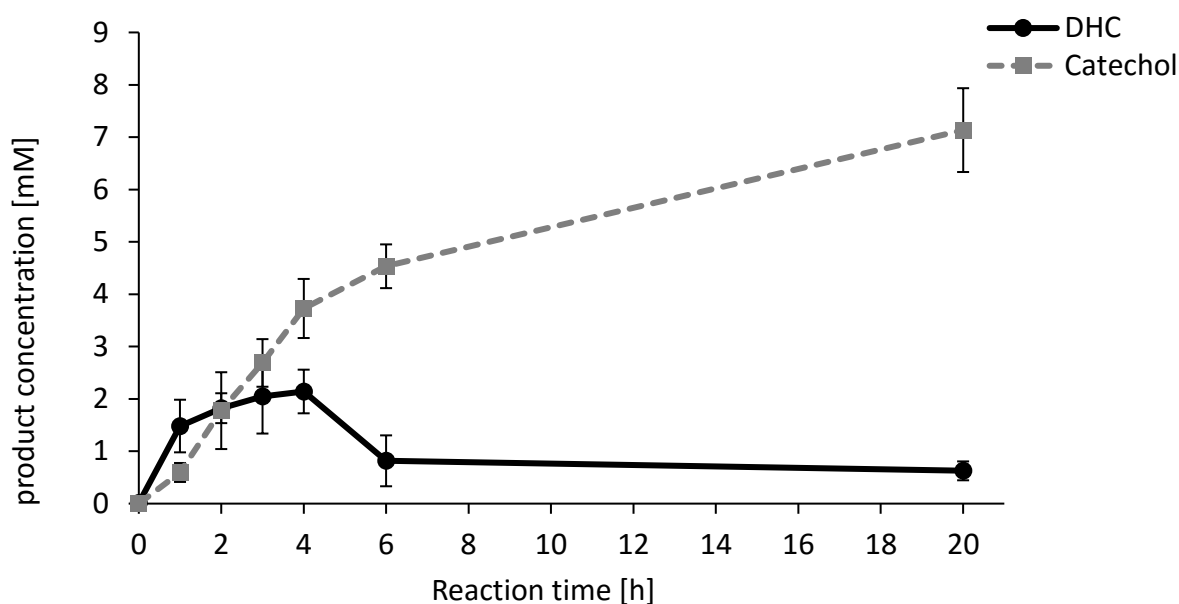
Figure 16: *E. coli* BW25113 pBAD18-TDO catalyzed conversion of 10 mM benzene **25** at different reaction times. The *cis*-dihydrocatechol **26** (DHC) concentration is depicted in black, and the catechol **41** concentration in dotted-grey. Protein expression for the whole cell biotransformations was performed in advance at 25°C for 20 h. Each of the measurement points represent the mean value of three biological triplicates and their corresponding standard deviations. Taken from Wissner and colleagues.[156]

It was also shown that with lower expression temperatures of 20°C, instead of 25°C, less **41** was formed at each measured time point (Figure 17). The influence of the expression temperature regarding **41** generation, suggested that a biocatalytic degradation of product **26** was taking place.
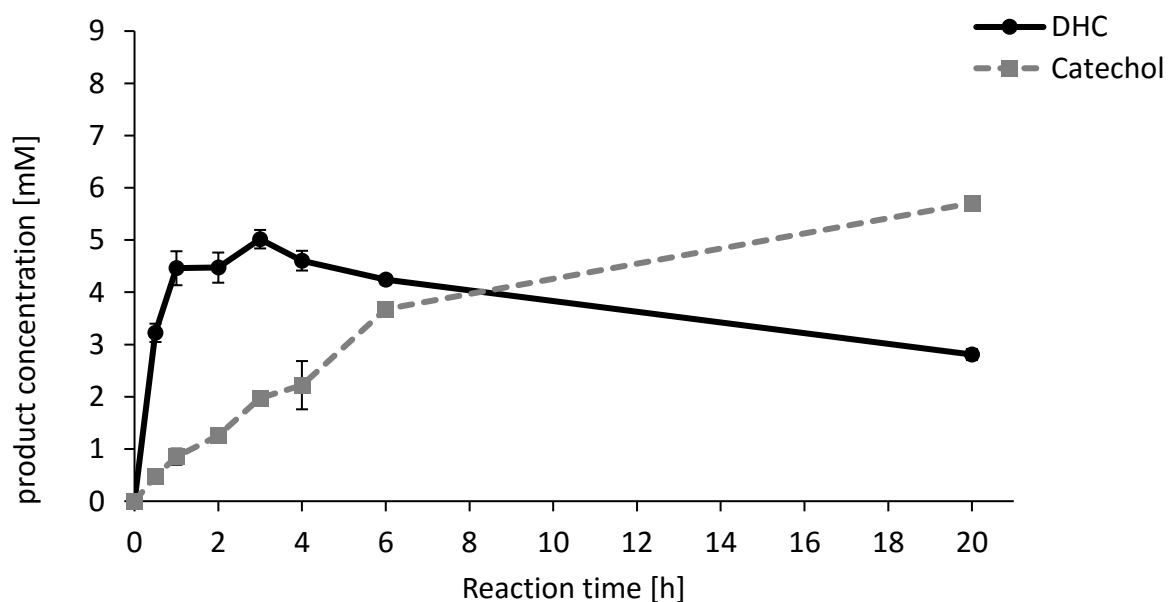
Figure 17: *E. coli* BW25113 pBAD18-TDO catalyzed conversion of 10 mM benzene **25** at different reaction times. The *cis*-dihydrocatechol **26** (DHC) concentration is depicted in black, and the catechol **41** concentration in dotted-grey. Protein expression for the whole cell biotransformations was performed in advance at 25°C for 20 h. Each of the measurement points represent the mean value of three biological triplicates and their corresponding standard deviations. Taken from Wissner and colleagues.[156]

Under these premises, an enhanced reaction-setup with a substrate feeding approach was established. This approach allowed to perform semi-preparative biotransformations in a volume of 100 mL, resulting in an isolated yield of 64 mg (14%) of **26**. In order to verify the hints regarding the biocatalytic driven dehydration of **26** to **41**, a basic experiment exploring the direct degradation of *cis*-dihydrodiendiol species was performed. Simple addition of biosynthesized **26** to the two different non-recombinant *E. coli* strains BW25113 and JM109(DE3) demonstrated that in both cases the cellular enzyme machinery was involved in the degradation of **26** to **41**. To fully confirm these findings, the same experiment was performed with heat-shocked lysates (95°C for 10 min), instead of whole cells. As expected, in both cases, the degradation of **26** into **41** was abolished completely, fully validating that an enzymatic reaction in *E. coli* and not a chemical driven degradation was responsible for the conversion of **26** into **41**. To verify whether the enzymatic dehydrogenation was selective for compound **26** or also occurred for other *cis*-dihydrodiendiols, *E. coli* BW25113 pBAD18-TDO-catalyzed biotrans-formations of toluene **11** were performed under the identical described reaction conditions. Results showed that the dehydrogenation took also place for the TDO-catalyzed product of **11**, *cis*-1,2-dihydro-3-methylcatechol **12** (Figure 18), although in much lower extent than for **26** (6% and 57% of the corresponding catechol formed entity after 20 h reaction, respectively). Thus, the addition of one methyl group to the substrate (from **25** to **11**) led to an almost 10-time lower catechol formation, suggesting that the involved enzyme in the *cis*-diol dehydrations preferred smaller substrates.
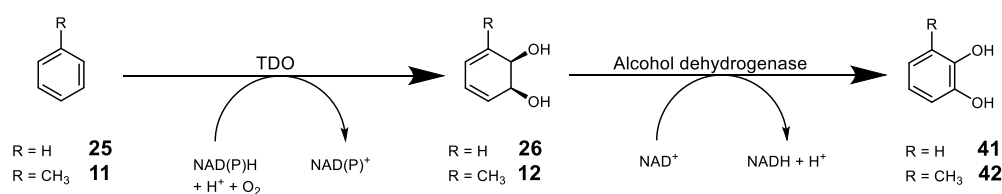
Figure 18: TDO-catalyzed conversion of benzene **25** and toluene **11** to *cis*-dihydrocatechol **26** and *cis*-1,2-dihydro-3-methylcatechol **12**, respectively. The formed products are degraded by a native alcohol dehydrogenase from *E. coli* to the undesired secondary products catechol **41** and 3-methylcatechol **42**, from **26** and **12**, respectively. Adapted from Wissner and colleagues.[156]

To further enhance the product yield of **26**, the search for the responsible dehydrogenating enzyme became an intrinsic task of this work, translating in its identification. In the native *P. putida* F1 strain, harboring the complete set of genes conforming the *tod* operon (Chapter 1.5.1, Fig 8B), *cis*-dihydrodiendiols are degraded by the downstream enzyme *cis*-toluene dihydrodiol dehydrogenase TodD.[95] This highly suggested that an alcohol dehydrogenase from the *E. coli* pool was capable of performing the same reaction, employing **26** as substrate. Initially, the NCBI Basic Local Alignment Search Tool[175] was used to perform a search of

candidate proteins in *E. coli* K-12 (the parent strain of BW25113), utilizing as query the protein sequence of TodD from *P. putida* F1. The output highlighted 3-phenylpropionate-dihydrodiol dehydrogenase HcaB (43% identity to TodD), as possible candidate. Indeed, the natural reaction of HcaB is the dehydrogenation of the substituted *cis*-dihydrocatechol-derivative *cis*-3-(3-carboxyethyl)-3,5-cyclohexadiene-1,2-diol to 3-(2,3-dihy-droxyphenyl)propio-nate.[176] Another possibility considered was the 2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase EntA of *E. coli*, which was described by Sakaitani and colleagues to perform the dehydrogenation of **26** to **41** *in vitro*.[177] To verify whether HcaB or EntA were the major responsible enzymes for the degradation of **26**, the corresponding single-gene knockouts of *E. coli* BW25113, Δ*hcaB* and Δ*entA,* were obtained from the KEIO collection and tested. Contrary to the expectations, neither BW25113 Δ*hcaB* nor Δ*entA* exhibited a major influence on **26** degradation, indicating that either another dehydrogenase or even multiple dehydrogenases were responsible for the degradation. To address the coverage of this question, a search for putative dehydrogenases was carried out using the online database EcoCyc,[178] which resulted in the identification of 156 confirmed and putative dehydrogenase and their corresponding genes in *E. coli* K-12 (the parent strain of BW25113). A medium-throughput system was then developed for a whole cell screening of all the identified dehydrogenase-deficient KEIO strains, in a 96 deep well plate configuration.[145] Thereby, the dehydrogenase deficient strains were cultivated in 96 deep well plates, and examined regarding their partial or total lack of capability to degrade **26**.[179] The robust and fast screening enabled the identification of the strain BW25113 Δ*gldA*, which degraded **26** only in minor amounts (<6% conversion). The gene *gldA* encodes the glycerol dehydrogenase GldA, which was identified as the main responsible for the degradation of **26**. Strikingly, GldA from BW25513 exhibits a 73% identity to the benzene *cis*-benzene dihydrodiol dehydrogenase BedD of *P. putida* ML2, but only a 15% identity to the initially searched *cis*-toluene dihydrodiol dehydrogenase TodD of *P. putida* F1 (Table 4).[156] The natural reaction of BedD is indeed the dehydrogenation of **26** to **41**. Thus, the high identity (73%) between BedD from *P. putida* ML2 and GldA from *E. coli* BW25113, suggested that the former was capable of performing the dehydrogenation of **26**. In addition, both enzymes share a similar substrate scope, for instance both are able to dehydrogenate the substrates glycerol and 1,2-propanediol.[180,181]

Summary of Research and Discussion

Table 4: Homology table of dehydrogenases. Identity was obtained by alignment of the corresponding proteins on UniProt.[182] Taken from Wissner and colleagues.[156]

	TodD (*P. putida*)	EntA ([b]*E. coli*)	HcaB ([b]*E. coli*)	BedD (*P. putida*)	GldA ([b]*E. coli*)	
TodD (*P. putida* F1)	100	23	43	17	15	
EntA ([b]*E. coli* K-12)		100	21	13	13	Identity
HcaB ([b]*E. coli* K-12)			100	10	12	
[a]BedD (*P. putida* ML2)				100	73	
GldA ([b]*E. coli* K-12)					100	

a; in literature also known as glyceroldehydrogenase GldA
b; K-12 is the direct parent strain of BW25113

The GldA-catalyzed dehydrogenation of *cis*-diol **26** was a novel finding, not previously reported in literature. Despite numerous publications utilizing recombinant *E. coli* strains for the production of the *cis*-diol **26**, not one of them described a secondary reaction of the formed product.[166,172,183] The influence of the growth and expression media is unlikely to be responsible for such discrepancy to literature, as rich media were used in both, the current study and two other literature reports that employed recombinant *E. coli* strains.[166,172] Furthermore, these publications may have been unable to detect **41** as byproduct due to their rather low formation of **26** (<8%), or simply ignored its formation, since they focused only on the generation of **26**. Nevertheless, a reference to the GldA-catalyzed degradation of compound **26** can be found in the work of Bagnéris and colleagues.[172] Here, they detected large amounts of **41** (33%) for the TDO-catalyzed conversion of **25**. However, they assumed that its formation was due to thermal decomposition during GC-MS analysis. Nevertheless, this supposition is highly questionable considering two points; firstly, the thermal decomposition product of **26** is phenol, since a dehydration occurs more readily than a dehydrogenation. Secondly, the TDO-catalyzed conversion of **11** performed by Bagnéris and coworkers yielded no trace of the corresponding 3-methylcatechol **42**, which should be detected, as thermal decomposition product. The absence of the degradation product **42** in the work of Bagnéris, can be explained with the findings in the current work, where **12** is degraded by *E. coli* to a much lesser degree than for **26**.

The discovery of GldA as main responsible for the degradation of **26**, allowed to eliminate the source of the undesirable secondary reaction. This enabled the development of a customized TDO-based platform aiming to enhance the production of **26**. Thus, chemically competent cells of *E. coli* BW25113 Δ*gldA* were generated and subsequently transformed with the pBAD18-TDO plasmid. Indeed, by using for biotransformation of **25** the recombinant *E. coli* BW25113 Δ*gldA* customized platform, harboring pBAD18-TDO, the secondary reaction leading to **41** was virtually completely abolished (Figure 19). This results fully verified that GldA was mainly responsible for the formation of **41**. Residual degradation of compound **26** may occur due to

other dehydrogenases, for instance EntA and HcaB, which seem to be capable of performing this reaction too, but in much lower amounts. In addition, the degradation of **12** during the biotransformation of **11**, could be reduced by half, by utilizing the enhanced BW25113 Δ*gldA* pBAD18-TDO platform (from 6% to 3% formation of the corresponding catechol after 20 h reaction), exhibiting the influence of GldA also for this secondary reaction.
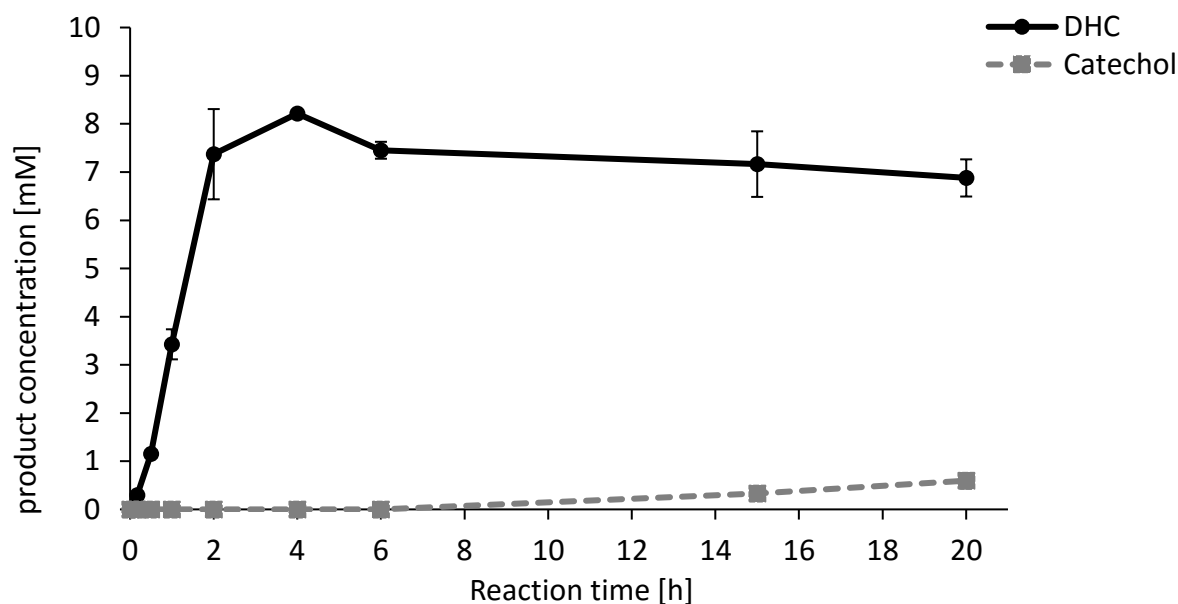
Figure 19: *E. coli* BW25113 Δ*gldA* pBAD18-TDO catalyzed conversion of 10 mM benzene **25** at different reaction times. The *cis*-dihydrocatechol **26** (DHC) concentration is depicted in black, and the catechol **41** concentration in dotted-grey. Protein expression for the whole cell biotransformations was performed in advance at 25°C for 20 h. Each of the measurement points represent the mean value of three biological triplicates and their corresponding standard deviations. Taken from Wissner and colleagues.[156]

Furthermore, lowering the expression temperature from 25°C to 20°C, resulted for biotransformations of substrate **25** in increased formation of product **26** and complete depletion of the undesired secondary reaction. Thus, after a reaction time of 4 h, 9.07 mM (90%) of product **26** could be prepared on analytical scale (Figure 20). The utilization of *E. coli* BW25113 Δ*gldA* pBAD18-TDO in a 100 mL semi-preparative biotransformation resulted in the isolation of 141 mg (31%) of the *cis*-diol **26**, after purification. Thus, the use of the selected *E. coli* strain BW25113 Δ*gldA* afforded a 2.2-fold higher production of **26**, in comparison to the strain BW25113, and additionally abolished the undesired secondary reaction yielding **41**.
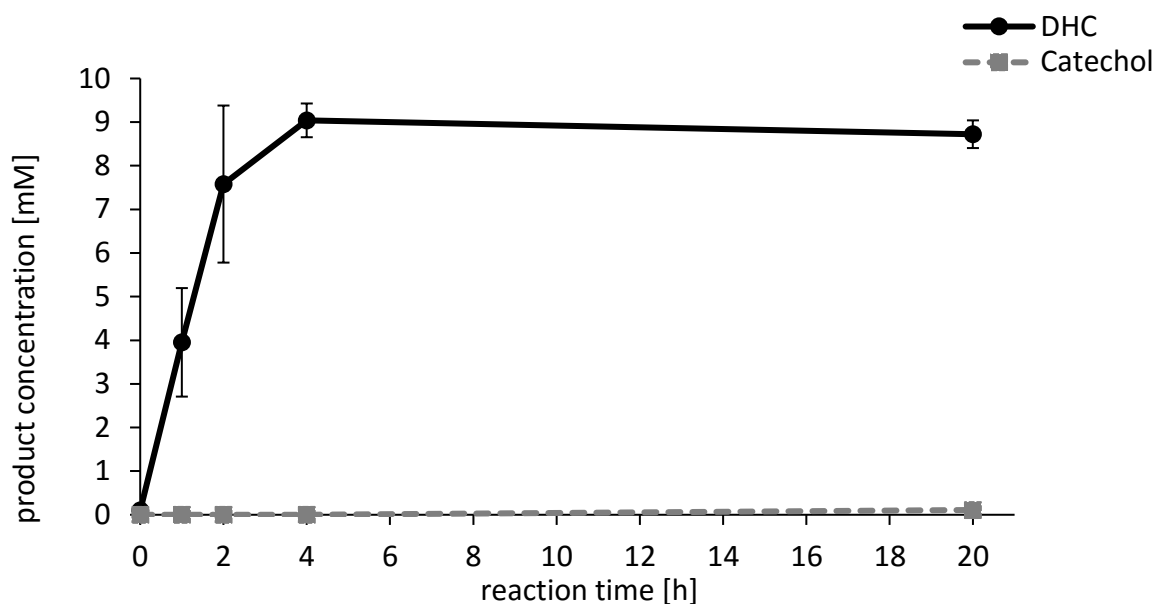
Figure 20: *E. coli* BW25113 Δ*gldA* pBAD18-TDO catalyzed conversion of 10 mM benzene **25** at different reaction times. The *cis*-dihydrocatechol **26** (DHC) concentration is depicted in black, and the catechol **41** concentration in dotted-grey. Protein expression for the whole cell biotransformations was performed in advance at 25°C for 20 h. Each of the measurement points represent the mean value of three biological triplicates and their corresponding standard deviations. Taken from Wissner and colleagues.[156]

Summarizing, a new efficient and attractive whole cell platform for the production of **26** was established. The TDO-based platform was developed based on the suitable characteristics of the pBAD18-TDO tunable system and on the strategic selection of the *E. coli* BW25113 Δ*gldA* cell factory. In comparison with previous literature, this platform exhibits multiple beneficial features, for instance shorter reaction times of only 6 h, the highest product formation of **26** so far reported, (90% in analytical scale), and the sole formation of the *cis*-dihydrodiendiol **26** with an advantageous isolated yield of 141 mg (31%).

3.2 Insights in research article II: An engineered toluene dioxygenase for a single step biocatalytical production of (-)-(1*S*,2*R*)-*cis*-1,2-dihydro-1,2-naphthalenediol

The successful establishment of the improved TDO platform, enabling the biosynthesis of substantial amounts of *cis*-dihydrocatechol **26,** settled the foundations for the next stages of the present doctoral work. Hence, the subsequent step was to investigate the catalytic behavior of the TDO-based platform towards larger and more challenging substrates. Therefore, the simplest bicyclic aromatic, naphthalene **20**, was selected as model substrate to evaluate the efficiency of the TDO driven conversion. The TDO-catalyzed conversion of **20** was previously described in literature, reporting the generation of (1*R*,2*S*)-1,2-dihydro-1,2-naphthalenediol **5a** as sole product in outstanding enantiomeric excess (>98%) and high yield (Figure 21).[84,184,185]
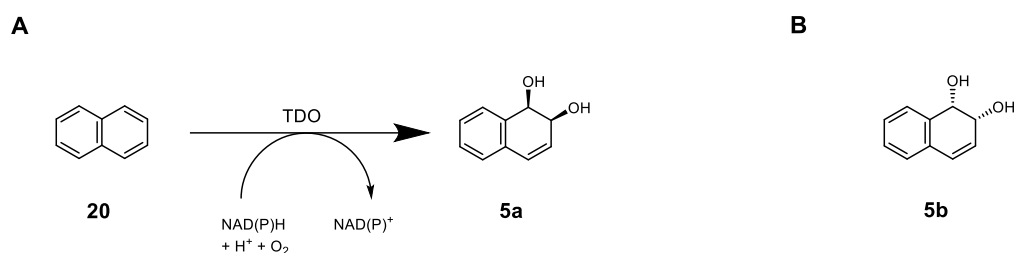
Figure 21: (**A**) TDO wild type catalyzed conversion of naphthalene **20** to (1*R*,2*S*)-1,2-dihydro-1,2-naphthalenediol **5a**. (**B**) The opposite enantiomer (1*S*,2*R*)-1,2-dihydro-1,2-naphthalenediol **5b** was not directly accessible, neither chemically nor biocatalytically, prior to the present work. Adapted from Wissner and coworkers.[157]

The valuable compound **5a** is a pharmaceutically relevant scaffold employed, for instance, in the synthesis of the amphetamine substitute (2*R*)-8-substituted-2-aminotetralins,[186] multiple bicyclic conduritol analogues,[187] and the cytotoxic compound (+)-goniodiol **6**.[34] Up to day, two chemical approaches for the synthesis of *cis*-1,2-dihydro-1,2-naphthalenediol are known, yielding either a racemic or an enantioenriched product mixture. The racemic generation of *cis*-1,2-dihydro-1,2-naphthalenediol requires a three step process, using naphthalene-1,4-dione as starting material (48% overall yield),[188] while enantiomerically enriched **5a** (90% *ee*) can be synthesized in four steps, utilizing rather expensive 2-vinylbenzaldehyde as substrate (29% overall yield).[189] In contrast, the RO-driven biosynthesis represents an attractive alternative to the purely chemical syntheses, since it involves solely a single-step reaction employing the readily available substrate naphthalene **20** to obtain enantiopure **5a**.

Since the establishment of the *E. coli* BW25113 pBAD18-TDO platform demonstrated the GldA-catalyzed dehydrogenation of *cis*-dihydrodiendiols,[156] initial experiments focused on questioning whether **5a** was also further converted. Thus, biotransformations were performed using BW25113 whole cells and commercially available racemic *cis*-1,2-dihydro-1,2-naphthalenediol as substrate. In this case, contrary to the observed degradation of product *cis*-dihydrocatechol **26**, the *E. coli* BW25113 strain was unable to degrade **5a**, indicating that the substrate scope of glycerol dehydrogenase GldA, responsible for the dehydrogenation of *cis*-diol **26** to catechol **41**,[156] does not include **5a**. These findings suggest that as substrate size increases (benzene, toluene, naphthalene), GldA-driven conversion decreases. Therefore, considering the absence of downstream degradation of **5a** in the host cells, further experiments were performed with *E. coli* BW25113 and not *E. coli* BW25113 Δ*gldA*. Thus, after the strategic selection of the *E. coli* BW25113 pBAD18-TDO platform, substrate **20** was converted. The performed biotransformations demonstrated the high efficiency of the platform to perform the expected reaction, since full conversion of 10 mM of substrate **20** was achieved in only 0.5 h, yielding enantiopure **5a** as sole product. The positive results prompted semi-preparative biotransformations in 100 mL scale, this time employing 20 mM of substrate **20**. Thus, after

20 h reaction and subsequent preparative high-pressure liquid chromatography (HPLC) purification, 287 mg (89%, >98% *ee*) of product **5a** were isolated.[157] The production of **5a** demonstrated the efficiency of the *E. coli* BW25113 pBAD18-TDO platform in comparison to the chemical synthesis and the enzymatic biosynthesis applying other RO systems reported in literature. Actually, the TDO-platform surpassed most of other approaches regarding the achieved product formation in percent and in reduced reaction time,[84,123,189,190] highlighting the feasibility of this novel approach for the synthesis of *cis*-diol **5a** in semi-preparative scale. Then, an additional milestone was established, pursuing the preparation of the opposite enantiomer, (1*S*,2*R*)-1,2-dihydro-1,2-naphthalenediol **5b**. It is worth to mention that to date there is no direct access to this interesting enantiomer, neither chemically nor biocatalytically. The state-of-the-art process for the preparation of **5b**, involves a three step chemical synthesis of racemic *cis*-1,2-dihydro-1,2-naphthalenediol, followed by a kinetic resolution with whole cells of *P. putida* strain 119 or NCIMB 11767 (19% overall yield).[191,192] Thereby, the *P. putida* strain is only capable of degrading the **5a** enantiomer, enabling the accumulation of **5b** in the cell suspension. Nevertheless, the atom economy of this process is not advantageous. There are numerous reports showing that both, NDO as well as TDO generate enantiopure **5a**.[84,184,187,190] However, no ROs have been reported to be able to produce the opposite enantiomer **5b** as main product. Nonetheless, there are a few reports attempting to reshape ROs in order to switch enantioselectivity. For instance, Parales and colleagues described the influence of the NDO active site variant F352V (corresponding to the F366 residue in TDO), enabling a modest shift in enantioselectivity, producing now **5a** with only 84% *ee*.[148] Indeed, this finding was of such importance that Parales and coworkers patented this NDO variant, which also displayed alterations in enantioselectivity for the substrates biphenyl and phenanthrene.[193] The corresponding mutation F366V in TDO was recently described by Vila and colleagues, who reported altered enantioselectivities for the substrates bromobenzene, styrene and indene.[153] In the case of the substrate indene, the enantioselectivity could not only be altered, but slightly switched in favor of the other enantiomer for the product (1*S*)-indenol (from +33 to -9). Considering these foundations, TDO position F366 was addressed as a target for site-directed mutagenesis, with the aim of switching the enantioselectivity towards **5b**. Initially, a sequence alignment of 671 ROs was performed with the 3DM database of Bio-Prodict (Nijmegen, NL). This approach enabled the design of a set of nine single-point variants at the TDO position F366. These variants, were generated and provided by the *PowerCart* project partner c-LEcta (Leipzig, DE), employing the newly developed pBAD18-TDO plasmid as parent template. Strikingly, biocatalytic explorations showed that eight out of nine variants were active and

displayed not only alterations in enantioselectivity, but a complete switch toward the pursued product **5b** (Figure 22). This phenomena can be explained by the crucial role of the phenylalanine residue at position 366, which is responsible for substrate coordination *via* hydrophobic interactions, as Vila and colleagues previously showed for the substrate bromobenzene and other substituted arenes by docking studies and computational modeling.[153,161]
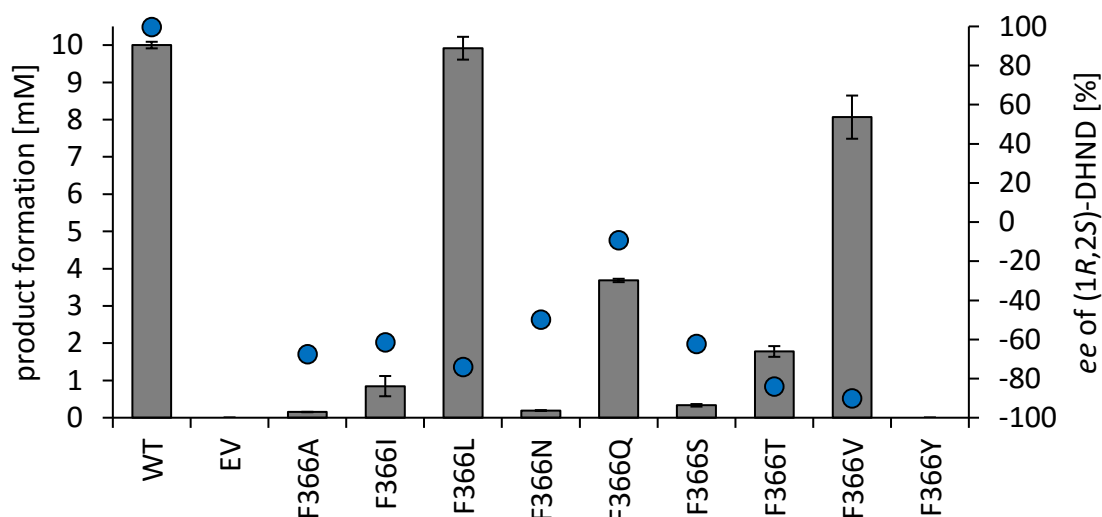
Figure 22: Product formation and enantiomeric excess of TDO variants at position F366, converting 10 mM naphthalene **20** in 20 h. The product formation is depicted in grey bars [mM] and the enantiomeric excess (*ee*) of the product (1*R*,2*S*)-dihydro-1,2-naphthalenediol **5a** as blue spheres [%]. WT stands for TDO wild type, and EV for empty vector, which was employed as negative control. Each of the shown points represent the mean value of three biological triplicates. The corresponding standard deviations for product formation are depicted. Adapted and customized for visualization from Wissner and coworkers.[157]

In this study, the most beneficial variant was F366V, which was not only able to switch enantioselectivity for **5a** from >98% *ee* to and outstanding -90% *ee*, but also exhibited a very good product formation of 8.07 mM (81%), when using 10 mM of substrate **20**. The final step was the semi-preparative biotransformation of 20 mM of **20** in 100 mL scale, utilizing the reshaped *E. coli* BW25113 pBAD18-TDO$_{F366V}$ platform, which yielded 101 mg (31%, 90% *ee*) **5b** after preparative HPLC purification. Hence, the envisaged milestone of generating a catalyst favoring the direct production of **5b** from substrate **20**, was successfully achieved after only one round of semi-rational mutagenesis, which allowed the identification of the TDO single point variant F366V as the most suitable biocatalyst for such purposes.

3.3 Insights in research article III: Semi-Rational Engineering of Toluene Dioxygenase from *Pseudomonas putida* F1 towards Oxyfunctionalization of Bicyclic Aromatics

Having demonstrated the efficiency of the pBAD18-TDO platform[156] and the achieved alterations in TDO enantioselectivity *via* semi-rational mutagenesis,[157] the next step was to take advantage of enzyme engineering for conversion of even more challenging bicyclic substrates harboring a heteroatom. According to the literature, TDO exhibits reduced activity with increasing substrate size. For instance, TDO-driven biotransformations of the well accepted flat substrate naphthalene **20** yield product concentrations three times higher than for the sterically more demanding biphenyl.[84] Furthermore, Boyd and coworkers obtained for the bicyclic substrate 2-chloroquinoline only a yield of 15% when performing biotransformations with recombinant *E. coli* JM109(pKST11) whole cells harboring TDO.[194] The addition of one or more heteroatoms into a biphenylic structure, decreases the activity substantially, yielding the corresponding *cis*-dihydroxylated product only in minor traces (1%).[195] Even bulkier substrates, such as the sterically demanding tricyclic azaphenanthrenes, are not accepted by TDO.

To circumvent such substrate scope limitations, the current work employed enzyme engineering to enhance the product formation as well as chemo-, regio-, and enantioselectivity of TDO towards bicyclic (hetero)aromatic substrates. Therefore, a semi-rational mutagenesis approach was envisaged focused on the active site, the putative substrate entrance channel, and key-positions previously reported in literature. The active site mutations were of particular interest, since a global mutagenesis study of all TDO active site residues was not performed before the present doctoral work. As mentioned in the previous publication, it was successfully shown that eight out of nine variants at position F366 exhibited activity, confirming that the 3DM database of Bio-Prodict (Nijmegen, NL) is a reliable and powerful tool to direct semi-rational mutagenesis.[157] Thus, based on the sequence alignment of 671 ROs, all active site variants with a conservation below 90% were chosen for site-directed mutagenesis. The highly conserved active site positions H222, H228, and D376, coordinating the catalytic iron,[102] as well as D219, crucial for electron transfer,[146] were not considered and remained untouched. The remaining 14 active site positions were mutated into amino acids, which according to the 3DM sequence alignment exhibited an occurrence of over 2% (Table 5). For active site positions L272 and I276, both located at a flexible loop structure, no consensus could be achieved. Thus, both were exchanged into the small amino acids glycine and alanine.

Summary of Research and Discussion

Table 5: List of all TDO active site positions with the corresponding amino acid exchanges generated semi-rationally in the course of this study. Taken from Wissner and colleagues.[158]

Position in TDO	Amino acid exchange
Q215	G, N
F216	A, G, I, L, N, Q, S, T, V, Y
[a]D219	-
M220	A, C, D, F, G, I, L, N, P, Q, S, T, V, W
[a]H222	-
A223	C, E, F, G, I, K, L, M, N, P, Q, S, T, V, W
[a]H228	-
G264	A, L, M, N, S
Y266	F, K, L, M, N, S
L272	A, G
I276	A, G
V309	A, F, G, H, N, S
H311	A, I, M, N, S, V, Y
L321	A, F, N, S
I324	A, F, G, M, N, S, T
F366	A, I, L, N, Q, S, T, V, Y
F372	A, G, L, N, Q, S, T, W, Y
[a]D376	-

a: Active site position D219 was not considered for mutation, owed to its high conservation (>99%). Active site position H222, H228, and D376 coordinating the catalytic iron exhibit a conservation level above 90% and thus, were not considered for mutagenesis.

In addition, the putative TDO substrate channel was calculated with the PyMOL software (version 2.4), employing the CAVER plugin.[196] To extend the putative substrate channel for bulky bicyclic substrates, six amino acids within a 5 Å proximity of the channel and outside the active site were mutated into the small amino acid alanine (Table 6).

Table 6: TDO positions located along the putative substrate entrance channel outside the active pocket. Amino acids within a radius of 5 Å of the simulated channel by the CAVER PyMOL plugin were replaced by alanine. Taken from Wissner and colleagues.[158]

Position in TDO	Amino acid exchange
G224	A
T225	A
L229	A
L245	A
P247	A
P248	A

Furthermore, a total of seven positions, either previously reported in literature,[107,149,153,197] or revealed by the 3DM Bio-Prodict database were chosen for mutagenesis (Table 7) and exchanged either according to literature or to the amino acid occurrence shown in the sequence alignment.

Table 7: The seven TDO hot-spot positions, either described in literature or suggested by the 3DM database of Bio-Prodict. Taken from Wissner and colleagues.[158]

Position in TDO	Amino acid exchange
F114	H, L, Y
T115	A, V
A212	T
A234	F, S
G323	A, C, Q
V340	A, F
T365	A, F, L, N

The complete set of 121 single-point variants located at 27 different positions, within the α-subunit of the TDO, were generated *via* site-directed mutagenesis by the cooperation partner c-LEcta (Leipzig, DE) as part of the *PowerCart* project. This set includes the previously described nine variants at the active site position F366.[157]

In order to evaluate the enzymatic activity of the generated mutant library, three bicyclic molecules were selected as substrates (Figure 23). First, the flat bicyclic substrate **20**, which is well accepted by TDO, generating the valuable synthon **5a**.[157] Following the successful switch in enantioselectivity with the TDO variant F366V, mutations that increase product formation were of particular interest. Second, the bulkier heteroaromatic compound 1,2,3,4-tetrahydro-quinoline **43**, a new substrate never reported before this study to be converted by TDO. In this case, the main targets were increased selectivity, improved product formation, and structural characterization of the TDO-catalyzed product(s). Third, the bulky substrate 2-phenylpyridine **46**, which is scarcely converted by TDO (1%).[195] The aim was to identify mutations increasing product formation and influencing regioselectivity.[158]

20   **43**   **46**

Substrate bulkiness

Conversion with TDO
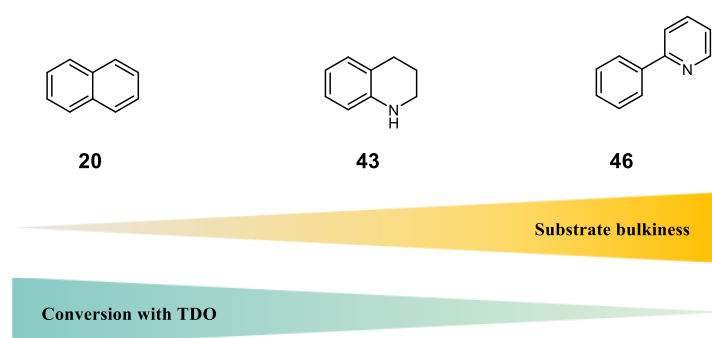
Figure 23: Bicyclic substrates employed to evaluate the enzymatic activity of the generated library. Naphthalene **20**, 1,2,3,4-tetrahydroquinoline **43**, and 2-phenylpyridine **46**. Substrate bulkiness (yellow gradient), increases from **20** to **46**, whereas the substrate conversion of TDO wild type (green gradient) displays an inverse correlation and decreases from **20** to **46**. Taken from Wissner and colleagues.[158]

The activity screening of the single point mutant library toward the three selected substrates was performed in 96 deep well plates, in biological triplicates, and analyzed by HPLC coupled with a diode-array detector (DAD) and an electrospray ionization mass spectrometer (ESI-MS). Variants, influencing the product formation or the chemo-, regio-, or enantioselectivity had to fulfill one or more of the following selection criteria to be considered as favorable;

1) Variants displaying over 200% total product formation in comparison to TDO wild type.
2) Variants displaying a switch in product distribution in favor to one of the side products and in addition, exhibiting in comparison to the wild type a product formation of at least 50%.
3) Variants generating selectively one product.

For full confirmation, all positive hits were verified by an enhanced reaction set-up in 96 deep well plates and additionally, their enantioselectivity was determined by chiral HPLC-DAD. Initial evaluation of the single point mutant library revealed variants with increased product formation at active site positions M220 and A223. Thus, a second mutant library was designed mainly based on variants of these two positions. The generated second mutant library comprised 55 double variants that were constructed and compared to their corresponding single point variants. Thereby, double variants had to be at least as active or as selective as the parent single variant, to be considered as beneficial and to be validated with the enhanced reaction set-up.

3.3.1   Conversion of naphthalene

Results showed that for substrate **20**, variants at active site positions M220 and A223 exhibited a tremendous influence in product formation (Figure 24). Especially variant TDO$_{A223V}$ was of great interest, since it enabled in comparison to the TDO wild type a 5.0-fold higher product formation, facilitating the production of 6.27 mM *cis*-diol **5a** with outstanding enantio-selectivity (>98%) in only 0.5 h. None of the double variants were able to surpass the product formation of TDO$_{A223V}$.
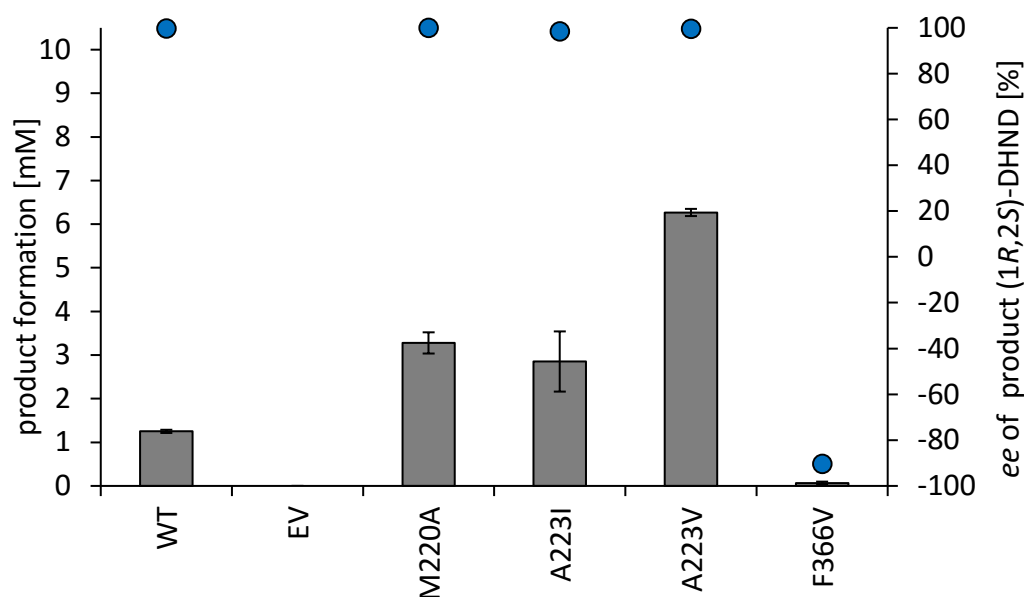
Figure 24: Product formation and enantiomeric excess of advantageous TDO variants, converting 10 mM naphthalene **20** in 0.5 h. The product formation is depicted in grey bars [mM] and the enantiomeric excess (*ee*) of the product (1*R*,2*S*)-dihydro-1,2-naphthalenediol **5a** as blue spheres [%]. WT stands for TDO wild type, and EV for empty vector, employed as negative control. Each of the measurement points represent the mean value of three biological triplicates. The corresponding standard deviation for product formation is depicted. Adapted and customized for visualization from Wissner and colleagues.[158]

As previously reported, TDO$_{F366V}$ generated the opposite enantiomer **5b** (90% *ee*). However, under the tested reaction conditions the observed conversion was much lower (0.06 mM).[157] Neither the double variant TDO$_{F216A\_F366V}$ nor TDO$_{M220A\_F366V}$ exhibited improvement in enantioselectivity (-70% *ee* and 13% *ee*, respectively) towards **5b**. None of the tested 176 variants influenced the chemo- or regioselectivity, since all active variants generated *cis*-dihydro-1,2-naphthalenediol as sole product.

3.3.2 Conversion of 1,2,3,4-tetrahydroquinoline

Conversion of the novel substrate **43** by TDO wild type yielded the monohydroxylated chiral product (*R*)-1,2,3,4-tetrahydroquinoline-4-ol **44** (83% *ee*) and the dehydrogenation product quinoline **45** (Figure 25). Interestingly, no *cis*-dihydroxylation of the aromatic moiety occurred. Such findings are in agreement with the TDO-catalyzed conversion of the similar molecule tetralin to the monohydroxylated product (*R*)-1-tetralol.[121] Thus, employing benzo-cycloalkenes and their heteroaromatic derivatives as substrates, α-allylic monohydroxylation occurs more readily than *cis*-dihydroxylation.

As expected, the substrate conversion was reduced in comparison to the flat substrate naphthalene **20**. While substrate **20** was converted completely within 20 h (10 mM), substrate **43** was only converted up to 30% (2.96 mM).
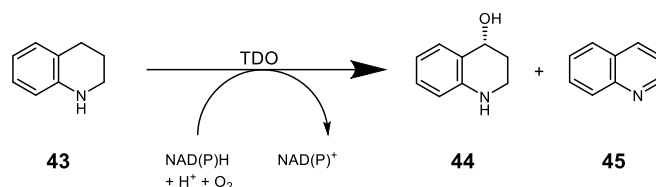
Figure 25: TDO wild type-catalyzed conversion of 1,2,3,4-tetrahydroquinoline **43** to the main product (*R*)-1,2,3,4-tetrahydroquinolin-4-ol **44** and the side product quinoline **45**. Adapted from Wissner and colleagues.[158]

The screening of the single point mutant library with substrate **43** derived in the master thesis performed by Jona T. Schelle.[198] These investigations highlighted active site positions A223 and L321, as well as position F114, located near to the ferredoxin binding site, to have a positive influence on product formation (Figure 26). Interestingly, the majority of variants at position M220 favored the production of the dehydrogenation product **45**, though at the cost of product formation. The previously described variant TDO$_{F366V}$ exhibited no influence on regioselectivity, but interestingly switched the chemoselectivity, producing now compound **45** in small quantities as main product. In addition, TDO$_{M220A}$ and TDO$_{L321S}$ generated **45** as main product, while exhibiting a reduction in total product formation.

Despite the interesting reaction itself, the dehydrogenation of substrate **43** to **45** is not a synthetic useful route. The TDO-catalyzed desaturation of a bicyclic compound is indeed not new, and was previously shown for the substrate 1,2-dihydronaphthalene,[123,199] but not specifically for **43**. Surprisingly, product **45** generated in this study is not further *cis*-dihydroxylated as previously reported by Boyd and colleagues.[159,200] They reported that compound **45** is accepted as substrate by TDO, yielding the dihydroxylated compound *cis*-5,6-dihydro-quinoline-5,6-diol as main product. Nevertheless, in biotransformations of **43** performed in this study no traces of the dihydroxylated product were detected, neither by employing HPLC-DAD, nor with more sensitive analytics such as HPLC-ESI-MS. An explanation could be the poor acceptance of compound **45** as substrate by TDO.[200] Moreover, rather low concentrations of **45** that could be employed as substrate were obtained (0.22 mM and 1.55 mM, for the wild type and for TDO$_{L321A}$, respectively). Another possibility, but one that has been discarded, is that compound **45** is a degradation product of the monohydroxylated compound **44**. Degradation of **44** would likely involve its dehydration (spontaneous or TDO catalyzed) to 7,8-dihydro-quinoline, followed by a TDO catalyzed dehydrogenation step to **45**. However, this degradation can be excluded, due to sole production of compound **44** by TDO$_{F114H\_A223T}$. Other factors hampering the formation of the dihydroxylated product might involve the genetic background, unforeseen downstream reactions, or operational factors such as, pH, reaction time and reaction temperature.

A

B
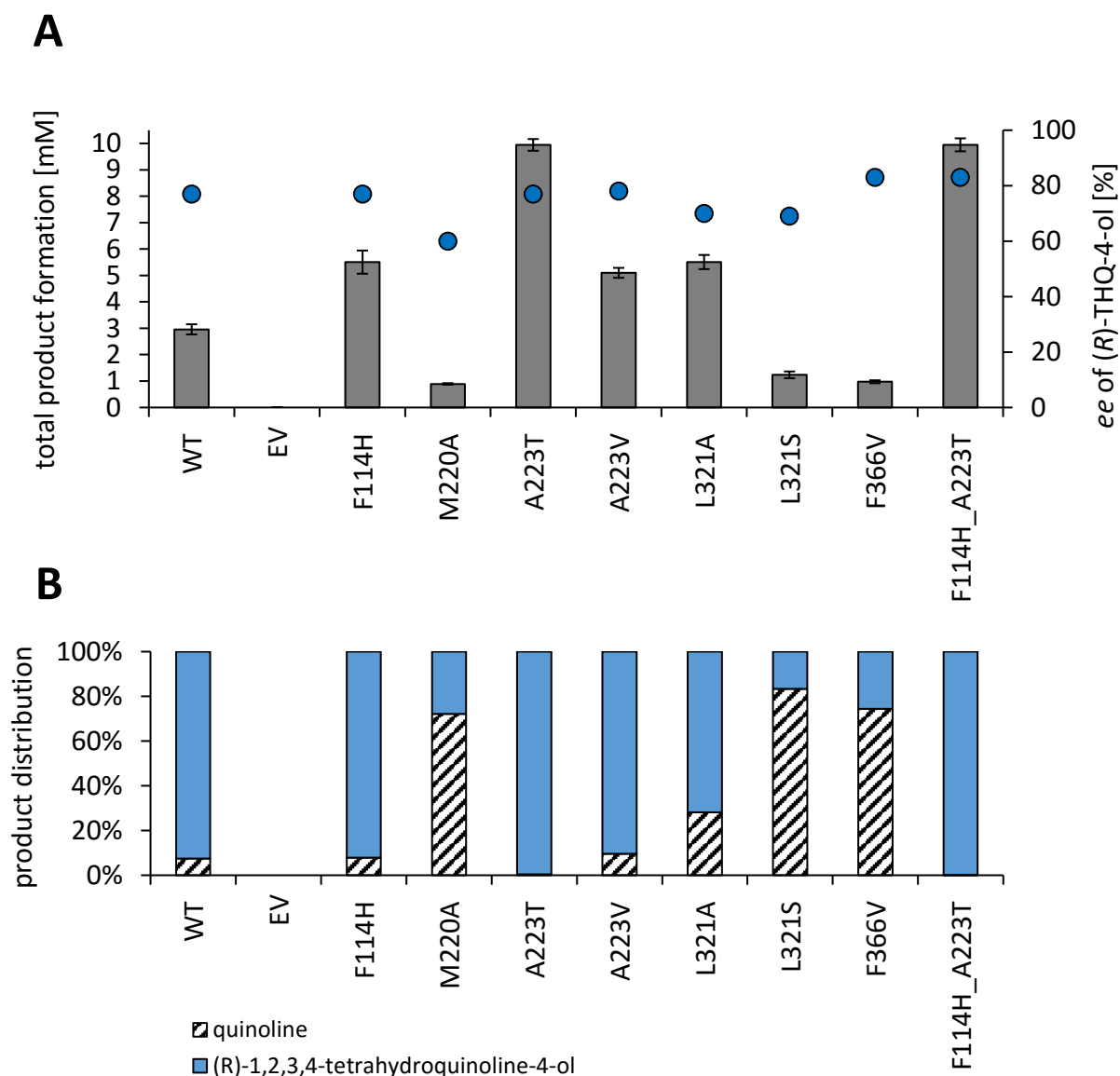
Figure 26: (**A**) Product formation and enantiomeric excess of beneficial TDO variants, converting 10 mM 1,2,3,4-tetrahydroquinoline **43** in 20 h. The total product formation is depicted in grey bars [mM] and the enantiomeric excess (*ee*) of the product (*R*)-1,2,3,4-tetrahydroquinoline-4-ol **44** ((*R*)-THQ-4-ol) as blue spheres [%]. WT stands for TDO wild type, and EV for empty vector, employed as negative control. Each of the measurement points represents the mean value of three biological triplicates. The corresponding standard deviation for product formation is depicted. (**B**) Product distribution [%] of the corresponding TDO variants. (*R*)-1,2,3,4-Tetrahydroquinoline-4-ol **44** is depicted in blue and quinoline **45** in striped-pattern. Adapted and customized for visualization from Wissner and colleagues.[157]

Double variant TDO_{F114H_A223T} enabled the complete conversion of 10 mM substrate **43**, yielding the pharmaceutically interesting monohydroxylated synthon **44** with enhanced enantioselectivity (83% *ee*), as sole product.[201,202] The enantiomeric excess could be further improved to 95% *ee*, by utilizing as reaction solvent 0.1 M tris(hydroxymethyl)aminomethane buffer (pH 8.5) instead of the routinely employed potassium phosphate buffer (pH 7.4). To date, only two routes can synthesize enantiopure compound **44**. The first route was described in 2014 by Zheng and colleagues, using *Pseudomonas plecoglossicida* ZMU-T02 whole cells to convert

substrate **43** to enantiopure **44** (82%, 99% *ee*).[203] A drawback of this biocatalytic route, is the formation of the undesired achiral side product 1,2,3,4-tetrahydro-2(1H)-quinolinone (16%). Additionally, compared to the TDO-catalyzed approach presented in the current study, Zheng and coworkers applied a rather modest substrate concentration of 2 mM. The second route was recently reported by Yin and colleagues, which utilized 1,2,3,4-tetrahydro-4-quinolinone as substrate.[204] The purely chemical route employed asymmetric hydrogenation and yields the monohydroxylated compound **44** as sole product with outstanding yield (99%, 98% *ee*). Nevertheless, the drawback of this route is clearly the special reaction conditions, such as oxygen free working conditions and the use of hydrogen atmosphere under high pressure (50 atm). In contrast, TDO_{F114H_A223T} is an attractive alternative for the selective production of compound **44** under environmentally friendly conditions.

In order to biosynthesize the two products, semi-preparative biotransformations were performed. TDO_{F114H_A223T} was employed for the generation of monohydroxylated compound **44**, yielding 106 mg (71%, 94% *ee*), whereas TDO_{F366V} was utilized for the generation of the dehydrogenation product **45**, yielding 20 mg (16%). Thereby, a new approach was applied for the semi-preparative biotransformations. Instead of performing one 100 mL one-pot reaction,[156,157] 200 simultaneously driven biotransformations, each 0.5 mL, in 96 deep well plates were performed and combined after reaction completion to yield a total volume of 100 mL. Though this approach is not considered to be an upscaling, it exhibited two major benefits in comparison to the one-pot reaction; First, higher product formation. The semi-preparative set-up in 96 deep well plates proved to be more consistent compared to semi-preparative 100 mL one-pot reactions. For instance, the semi-preparative one-pot bio-transformations of benzene **25**[156] and naphthalene **20**[157] exhibited in comparison to the corresponding analytical scale a 1.5-2.5-fold decrease in product formation. Thus, the 200 simultaneous biotransformations set-up in 96 deep well plates offered not only a more reliable system, but also resulted in a higher product formation. The possible reasons in addition to different growth and expression conditions in Erlenmeyer flasks compared to 96 deep well plates, are for instance the working factors, such as dissolved oxygen and surface to volume ratio. Second, more economical. Only 96 mL TB-media is required for the generation of an average of 5 g_{cww} cell pellets harboring active TDO in 96-deep well plates, whereas one 2 L Erlenmeyer flask containing 500 mL TB media only provides an average of 4 g_{cww} cell pellets harboring active TDO. Thus, 200 simultaneous biotransformations are quite convenient in cost-benefit terms, since grow media consumption is reduced substantially.

3.3.3   Conversion of 2-phenylpyridine

In agreement with literature, the TDO-catalyzed conversion of 2-phenylpyridine **46** yielded the *cis*-dihydroxylated compound (1*S*,2*R*)-3-(pyridin-2-yl)cyclohexa-3,5-diene-1,2-diol **47**, in minor amounts (0.50 mM, 5%) (Figure 27). One application of this interesting *cis*-diol, is the generation of the bioactive bridged bicyclic compounds *via* Diels-Alder reaction.[205] Additionally, TDO catalyzed also the formation of the literature unknown side product 2-phenylpyridine-3-ol **48** (0.11 mM, 1%).[195] The finding of the latter compound was of great interest, since the RO-catalyzed hydroxylation of heteroaromatic moieties in biphenyl derivatives was not reported before the present doctoral thesis. Its formation involves most likely the formation of an unstable *cis*-dihydrodiendiol, which spontaneously rearomatize under the elimination of one water molecule, as described for the TDO-catalyzed hydroxylation of substituted pyridines.[150,206]
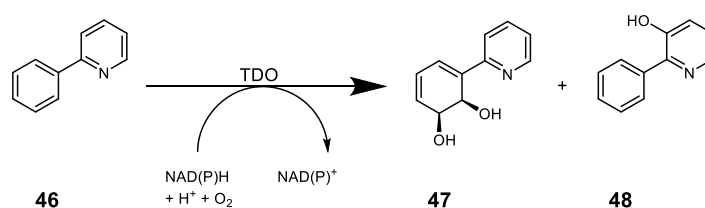
Figure 27: TDO-catalyzed conversion of 2-phenylpyridine **46** to the main product (1*S*,2*R*)-3-(pyridin-2-yl)cyclohexa-3,5-diene-1,2-diol **47** and the side product 2-phenylpyridine-3-ol **48**. Adapted from Wissner and colleagues.[158]

Screening of the single and double mutant libraries with substrate **46**, highlighted variants at the active site positions M220, A223 and V309, regarding increased product formation (Figure 28). As a result, double variant TDO_{M220A_V309G} not only enabled a 15-fold higher product formation, in comparison to the wild type, but also shifted the selectivity towards the formation of the *cis*-diol (8.51 mM **47** and 0.64 mM **48**). In contrast, variants TDO_{F216A} and TDO_{F366V} exhibited a decreased product formation, but generated instead the monohydroxylated compound **48** as main product, in the case of the latter variant even exclusively. Interestingly, none of the described variants altered enantioselectivity, thus, all produced the *cis*-diol **47** in outstand enantiomeric excess (>98%).

A

B
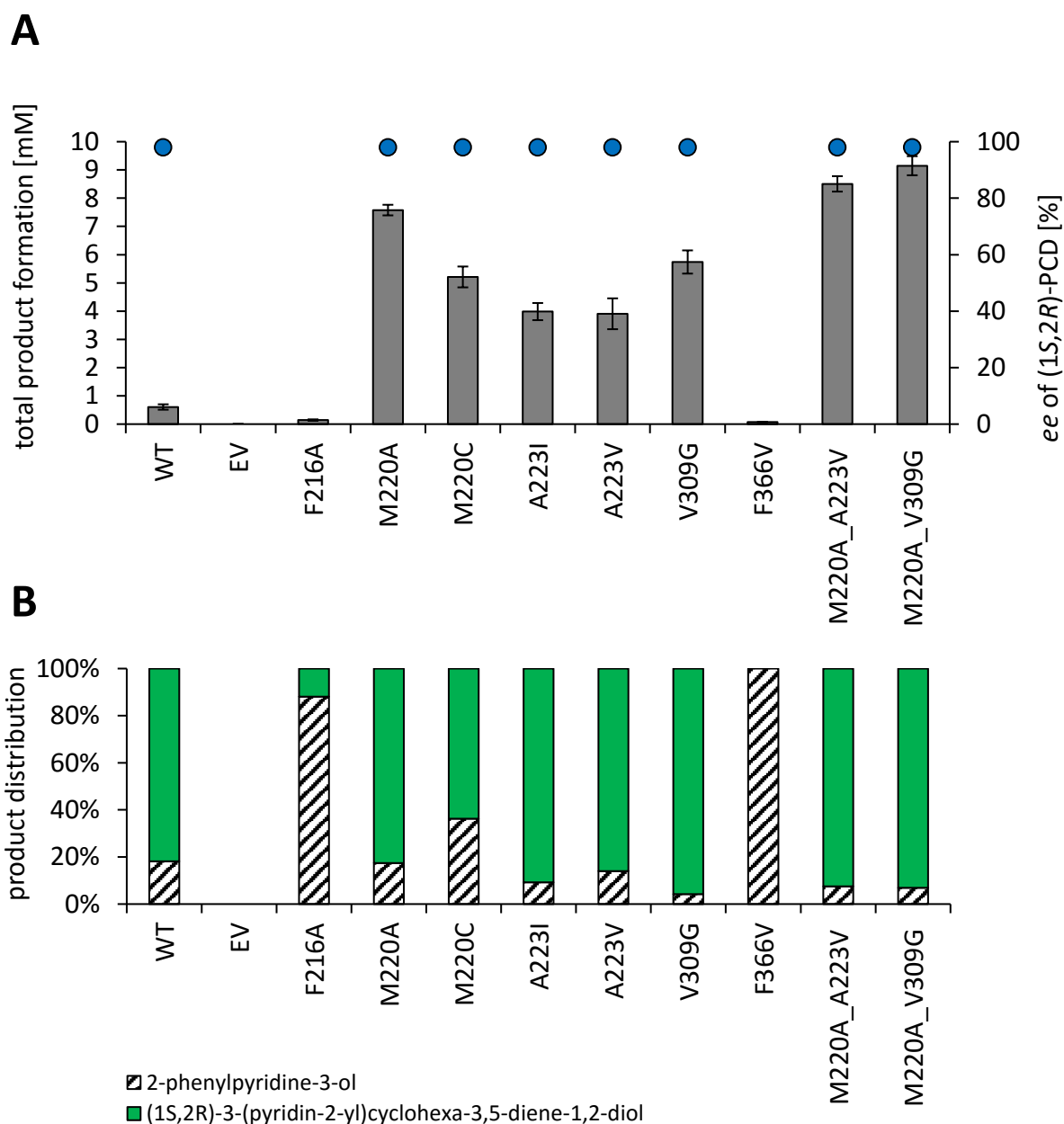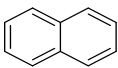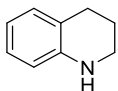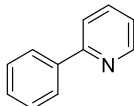
Figure 28: (**A**) Product formation and enantiomeric excess of advantageous TDO variants, converting 10 mM 2-phenylpyridine **46** in 20 h. The total product formation is depicted in grey bars [mM] and the enantiomeric excess (*ee*) of the product (1*S*,2*R*)-3-(pyridin-2-yl)cyclohexa-3,5-diene-1,2-diol **47** ((1*S*,2*R*)-PCD) as blue spheres [%]. WT stands for TDO wild type, and EV for empty vector, employed as negative control. Each of the measurement points represent the mean value of three biological triplicates. The corresponding standard deviation for product formation is depicted. (**B**) Product distribution [%] of the corresponding TDO variants. (1*S*,2*R*)-3-(Pyridin-2-yl)cyclohexa-3,5-diene-1,2-diol **47** is depicted in green, 2-phenylpyridine-3-ol **48** in striped-pattern. Adapted and customized for visualization from Wissner and coworkers.[157]

For isolation and characterization of the two products, semi-preparative biotransformations in 96 deep well plates (total volume of 100 mL) with variant TDO$_{M220A\_V309G}$ were performed, yielding 114 mg (60%, >98% *ee*) of *cis*-diol **47** and 6 mg (4%) of the monohydroxylated compound **48**.

51

Summary of Research and Discussion

Thus, these results demonstrate that the initial aim to enhance the TDO-catalyzed conversion of bicyclic aromatics *via* enzyme engineering was achieved. This translates in a substantial improvement in conversion of the substrates **4**3 and **46**, yielding the attractive chiral products **44** and **47**, respectively, with high chemo-, regio-, and enantioselectivity. Furthermore, seven TDO positions having a fundamental influence on product formation, chemo-, regio-, or enantioselectivity for at least one of the three bicyclic substrates were identified (Table 8).

Table 8: TDO positions with prominent influence on chemo-, regio-, and enantioselectivity and positions exhibiting, in comparison to the wild type, an increased product formation for the bicyclic substrates; naphthalene **20**, 1,2,3,4-tetrahydroquinoline **43**, and 2-phenylpyridine **46**. Adapted from Wissner and colleagues.[157]

	20	43	46
Product formation	M220, A223	F114, A223, L321	M220, A223, V309
Chemoselectivity	-	M220, A223, L321, F366	-
Regioselectivity	-	-	F216, M220, V309, F366
Enantioselectivity	F366	-	-

It is worth to mention that six of the seven key-positions (F216, M220, A223, L321, V309 and F366) are located in the active site (Figure 29), while interestingly, the seventh position F114 is located near the putative ferredoxin binding site, suggesting a crucial role in ferredoxin fixation.
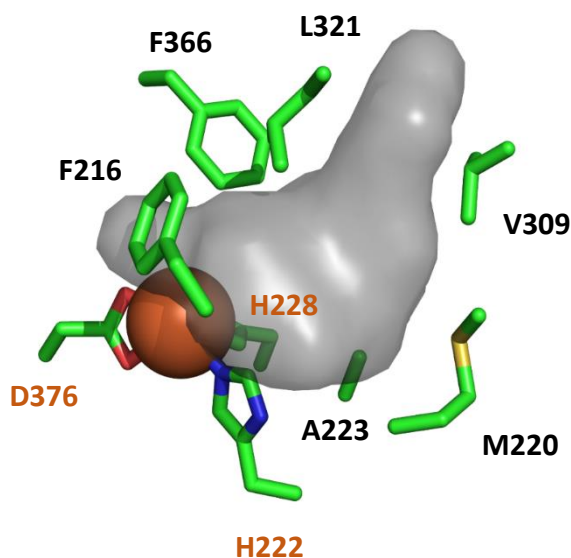
Figure 29: Close up of the active site of the TDO wild type. The active pocket, depicted in grey, is surrounded by the six highlighted key-positions F216, M220, A223, L321, V309 and F366. To facilitate orientation to the active pocket, active site residues H222, H228 and D376, which coordinate the catalytic iron (orange sphere), are included (orange). The model was generated with the software PyMOL (version 2.4). Taken from Wissner and coworkers.[157]

Summary of Research and Discussion

Similar tendencies were observed by comparing these seven hot-spot positions with the corresponding ones to the homologous enzyme NDO described by Halder.[26,154] For instance, variants of the position F202, A206, and H295 (corresponding to F216, M220, and V309 in TDO, respectively) increased the product formation for monocyclic aromatic substrates. In addition, chemoselectivity was altered by NDO variants at positions A206, F352, and L307 (corresponding to M220, F366, and L321 in TDO, respectively) and regioselectivity by F352 (corresponding to F366 in TDO).

In the current work, the three active site positions M220, A223, and F366 were of special interest, since they exhibited for all of the bicyclic substrates a crucial influence. For instance mutation at position M220 into an alanine led for the substrates **20** and **46** to an increased product formation, and to an altered the chemoselectivity for the substrate **43** in favor to the dehydrogenation product **45**. The influence of this position was previously described in literature. Thus, by employing variant TDO$_{M220A}$, Beil and coworkers enabled the conversion of the substrate 1,2,4,5-tetrachlorobenzene which was not previously accepted by TDO wild type.[147] Such behavior was also reported for the TDO homologues NDO and CDO, where the corresponding variant A206V and M232A, respectively, exhibited an influence in product formation, chemo-, regio-, and enantioselectivity.[152,154] Therefore, mutation of this position in TDO or its homologues, into smaller amino acids, like alanine, valine or cysteine seems to enlarge the active site, which leads to higher product formation or altered selectivities.

Variants at position A223, for instance TDO$_{A223V}$, led for all three substrates to an increased product formation. In addition, variant A223T exhibited an influence in chemoselectivity, enabling the exclusive production of the valuable monohydroxylated product **44**. The central influence of TDO position A223 was so far not reported in literature, neither for TDO nor for other homologue ROs. Only one single publication reported an effect of the corresponding position V205 in the aniline dioxygenase from *Acinetobacter sp.* strain YAA, where a mutation into an alanine led to a decreased product formation, but expanded the substrate scope.[207]

TDO position F366 exhibited the most versatile influence of all tested positions. Especially, TDO$_{F366V}$ was of great interest, since it exhibited for each of the three substrates another effect. As previously reported, TDO$_{F366V}$ switched completely the enantiomeric excess for substrate **20**.[157] Besides, by employing this variant for biotransformations of substrate **43**, a switch in chemoselectivity, favoring now the production of the dehydrogenation product **45**, could be achieved. Furthermore, TDO$_{F366V}$ enabled for the bulky substrate **46** a complete switch in regioselectivity, leading to the generation of the monohydroxylated compound **48** as sole product. Nevertheless, in comparison to the wild type, all TDO variants at position 366

exhibited a reduced product formation. The versatile effect of TDO variants at position 366 was described in literature, not only for TDO, but also for the RO homologues NDO and biphenyl dioxygenase (BPDO) from *Pseudomonas pseudoalcaligenes* KF707.[148,153,208]

Interestingly, all of the 14 non-conserved active site positions (Table 5) could be exchanged against other amino acids, and exhibited at least for one of the generated variants, still residual activity for one of the tested substrates. Thus, none of these amino acids in the active site of TDO are irreplaceable for hydroxylation reactions. For instance, 87 of the 98 generated active site variants exhibited activity towards substrate **43**. This shows that active site positions of TDO are quite tolerant to cope with mutations and can easily be replaced, without complete enzyme inactivation.

Exchange of the six substrate channel positions into the small amino acid alanine (Table 6) did not increased product formation. Nevertheless, all six alanine variants were able to convert each of the three substrates, exhibiting, depending on the substrate, equal or less product formation compared to TDO wild type. Therefore, the supposed broadening of the substrate channel did not lead to an increase in product formation, indicating that the active pocket is the bottle neck for conversion of larger substrates. Nevertheless, based on the residual activity of the alanine variants, none of these substrate channel positions proved to be crucial for enzyme activity or substrate binding.

Recently, Preston-Herrera and colleagues revealed an alteration in substrate selectivity for TDO variants at active site position L272.[209] Additionally, Li and coworkers described variant S283M in BPDO from *Burkholderia xenovorans* LB400 (corresponding to L272 in TDO), which exhibited an influence on substrate specificity and regioselectivity.[210] Neither TDO$_{L272A}$ nor TDO$_{L272G}$ generated in this study altered the selectivity for bicyclic substrates. Thus, the generation and characterization of TDO$_{L272M}$ and TDO$_{L272V}$ could provide interesting results regarding the selectivity for bicyclic substrates.

Mohammadi and colleagues reported for the single mutant T335A in BPDO from *Burkholderia xenovorans* LB400 (corresponding to G323 in TDO), an increased conversion of dibenzofuran.[197] However, all TDO variants at position G323 generated in the current study exhibited a decreased product formation for all three substrates. Thus, the trend of the former publication could not be verified for TDO with the substrate panel presented here.

Vila and colleagues described the major influence of TDO$_{I324F}$ and TDO$_{T365N}$ on chemo- and regioselectivity.[153] The generated I324 variants in the present study altered slightly the chemoselectivity in favor to the monohydroxylation of substrate **43**, but exhibited no change in selectivity for substrate **46**. Surprisingly, none of the generated T365 were able to change the

chemo- or regioselectivity substantially. However, variants at position I324 as well as T365 showed, in comparison to the wild type, a decreased conversion of all three substrates, which hampered the exact determination of the product ratios and therefore selectivities.

Hence, the insights about research article III aimed to provide the highlights achieved by utilizing the advantages of enzyme engineering for the conversion of even more challenging substrates, including heteroaromatic bicycles.

4   CONCLUSION AND OUTLOOK

Rieske non-heme iron dioxygenases (ROs) are powerful catalysts for the selective *cis*-dihydroxylation of aromatic compounds. The generated *cis*-dihydrodiendiols are versatile synthons employed for the generation of natural products, pharmaceuticals and fine chemicals.[1,24] The application of such biocatalytically generated *cis*-diols often significantly shortens synthetic routes for the production of several compounds, in comparison to the purely chemical ones. One of the most studied and widely characterized ROs in literature is toluene dioxygenase (TDO) from *P. putida* F1. This attractive enzyme combines a large substrate scope with outstanding product formation and enantioselectivity. Nevertheless, in the last decades, in spite of the frequent utilization of recombinant *E. coli* cells harboring TDO, neither the plasmid system nor the *E. coli* strain was optimized.[99,166,167] Moreover, though some bicyclic aromatics are within the substrate scope of TDO, product formation is substantially reduced in comparison to monocyclic aromatics, hampering the actual application of TDO in demanded synthetic approaches.[84]

In the course of the present thesis, a robust and enhanced TDO-based platform for the selective *cis*-dihydroxylation of aromatics was generated and designated as *E. coli* BW25113 pBAD18-TDO.[156] The unforeseen dehydrogenation of smaller monocyclic *cis*-dihydrodiendiols to their corresponding catechols, performed by *E. coli*, was circumvent by identifying the enzyme glycerol dehydrogenase (GldA) as responsible for the downstream degradation.[179] Based on this knowledge, the customized platform *E. coli* BW25113 Δ*gldA* pBAD18-TDO was established, allowing to practically abolish the unwanted secondary reaction, enabling the formation of 9.0 mM (90%) of *cis*-dihydrocatechol **26** from the simple and accessible substrate benzene **25**, in analytical scale. These results settled the foundations to perform a semi-preparative biotransformation of **25** in 100 mL scale, resulting in the selective generation of 141 mg (31%) of the valuable product **26**. Furthermore, it was found that the substrate scope for the GldA-driven degradation only includes monocyclic *cis*-dihydrodiendiols, but not bicyclic ones. Therefore, the studies involving bicyclic aromatics in this doctoral work were performed with the platform BW25113 pBAD18-TDO. Thus, by applying this system, it was possible to enable the complete conversion of 10.0 mM naphthalene **20**, and to perform a 100 mL semi-preparative scale biotransformation, which resulted in 287 mg (89%) of enantiomerically pure (1*R*,2*S*)-1,2-dihydro-1,2-naphthalenediol **5a** (>98% *ee*).[157] Then, in order to confer a switch in product enantioselectivity, TDO single point variants at hot-spot position F366 were generated, revealing the striking influence of this position. By utilizing

TDO$_{F366V}$, the enantioselectivity was switched completely, producing now in analytical scale 8.07 mM (81%) of (1S,2R)-1,2-dihydro-1,2-naphthalenediol **5b** (90% ee). To prove the feasibility of synthetizing substantial amounts of **5b**, a semi-preparative biotransformation in 100 mL scale was performed, yielding 101 mg (31%, 90% ee) of this valuable enantiomer. It is worth to remark that before this doctoral thesis, the compound **5b** was not directly accessible, neither chemically, nor biocatalytically.

Furthermore, by evaluating a semi-rational designed TDO mutant library consisting out of 176 single and double variants, mainly addressing the active site, a substantially improved product formation was accomplished, as well as enhanced chemo-, regio-, and enantioselectivity towards bulky bicyclic (hetero)aromatics.[158] For instance, TDO$_{F114H\_A223T}$ was not only capable of converting 10 mM of 1,2,3,4-tetrahydroquinoline **43** completely, exhibiting in comparison to the wild type an 3.6-fold increased product formation, but also abolished the generation of the side product quinoline **45** and produced exclusively (R)-1,2,3,4-tetrahydro-quinoline-4-ol **44**. Double variant TDO$_{M220A\_V309G}$ enabled an astonishing 15.1-fold higher conversion of the substrate 2-phenylpyridine **46**, in comparison to the wild type, and in addition altered the regioselectivity towards the cis-dihydroxylation of the aromatic moiety, generating 8.51 mM (1S,2R)-3-(pyridin-2-yl)cyclohexa-3,5-diene-1,2-diol **47** and 0.64 mM 2-phenylpyridin-3-ol **48**. The establishment of a miniaturized semi-preparative set-up in 96 deep well plates, designed to perform simultaneous and reproducible biotransformations, enabled to handle semi-preparative reaction volumes of 100 mL in a practical and cost-effective way. By using this approach, 100 mL of combined TDO$_{F114H\_A223T}$ biotransformations of **43** yielded 106 mg (71%, 94% ee) of the monohydroxylated product **44**, while 100 mL of combined TDO$_{M220A\_V309G}$ biotransformations of **46** afforded 114 mg (60%, >98% ee) of enantiomerically pure cis-diol **47** and 6 mg (4%) of the monohydroxylated compound **48**. Thus, the initial aim to foster the product formation of bulky bicyclic (hetero)aromatics via enzyme engineering was achieved. Thereby, variants at the active site positions M220, A223, and F366 exhibited a substantial influence either in product formation or in chemo-, regio-, and enantioselectivity, for each of the tested substrates. For instance, in addition to the literature known hot-spot position M220, the results of the current thesis highlighted position A223 as key-position regarding boosting product formation. TDO$_{A223V}$ increased the conversion of all three substrates utilized, while TDO$_{A223T}$ enabled the selective production of chiral compound **44** with increased product formation. Till today, the positive influence of TDO position A223 on product formation was not described, neither for TDO, nor for other homologue ROs. Variants at the hot-spot position F366, especially TDO$_{F366V}$, proved to be highly versatile, since

in dependence of the substrate, it displayed a notable influence in chemo-, regio-, and even enantioselectivity.

A peculiar observation was that each of the 14 non-conserved active site residues addressed in this study, could be mutated without the loss of residual activity for at least one of the substrates. Thus, the active site of TDO seems to be highly tolerant against mutations.

To further exploit the great potential of the enhanced TDO platform, different strategies could be implemented. For instance, a combinatorial active site saturation mutagenesis of the three hot-spot positions M220, A223, and F366 could be highly interesting for the determination of synergistic effects on product formation and chemo-, regio-, and enantioselectivity.[211] Thereby, degenerated NNK (N; any nucleotide, K; G or T) are introduced at the positions of interest, creating a multitude of mutants. For instance, such approach would result in 8000 different mutants for the combinatorial saturation of three positions, with a total of 98164 clones in the screening pool to be tested, in order to achieve a 95% probability of mutant coverage. Alternatively, the codon redundancy and the screening effort can be reduced significantly if the 22c-trick is employed.[212] Hereby, a special mixture of three primers create a degeneration of 22 unique codons, coding for all 20 amino acids. However, this approach would still result in a large number of 31899 clones to screen, which would require a robust medium- or high-throughput screening system. In this sense, the indigo assay provides a suitable high-throughput colorimetric solid-phase assay to screen RO activity.[148] It has been shown that this reliable, cheap and easy to use system is not only applicable to the NDO system, but also for the *E. coli* BW25113 pBAD18-TDO platform.[145]

Another interesting study, focused on the alteration of enantioselectivity, could be the characterization of TDO$_{F366V}$ for the conversion of alkylated or halogenated naphthalene derivatives. Hence, of great interest is the generation and isolation of the opposite enantiomers, which to date are not directly accessible.[213]

Furthermore, during the last years, the RO-catalyzed selective hydroxylation of monoterpenes received increased attention. Especially the CDO-catalyzed conversion of (*R*)-limonene to (1*R*,5*S*)-carveol was the topic of multiple studies.[152,155,214] This valuable product is not only an important flavor and medicinal compound with a great market potential,[215] but also can be employed for the biocatalytic production of carvolactone, which is an interesting building block for bioactive or natural products and also as monomer for polymer production.[216] Nevertheless, up to the present day the TDO-catalyzed oxyfunctionalization of (*R*)-limonene and other monoterpenes was not described. Thus, the characterization of TDO and its variants regarding

the hydroxylation of monoterpenes could open the door for the generation of highly sought-after compounds.

A novel perspective was recently provided by Heinemann and colleagues, which reported the dramatic influence of loops located near the active site of CDO.[155] They described that introducing insertions, deletions, and point mutations in such loops, can affect enzyme activity as well as the regio- and enantioselectivity for several substrates. In line with this, it can be anticipated that the combination of TDO active site variants with such loop modulations could generate interesting results for a variety of substrates.

In summary, ROs efficiently catalyze the highly sought-after *cis*-dihydroxylation of aromatics, which until to date are still not accessible by chemical synthesis. As demonstrated in the course of this thesis, enzyme engineering is a powerful tool able to substantially improve the activity and selectivity of ROs, generating enhanced biocatalysts that are more suitable for industrial applications. Consequently, even after half a century of extensive studies, the field of RO-catalyzed reactions remains a fascinating area worthy of further exploration.

5 LITERATURE

[1]   R. E. Parales, S. M. Resnick, in *Biocatalysis in the Pharmaceutical and Biotechnology Industries*, CRC Press, **2006**, pp. 299–331.

[2]   J. Chapman, A. E. Ismail, C. Z. Dinu, *Catalysts* **2018**, *8*, 238.

[3]   S. I. Shoda, H. Uyama, J. I. Kadokawa, S. Kimura, S. Kobayashi, *Chem. Rev.* **2016**, *116*, 2307–2413.

[4]   B. M. Nestl, B. A. Nebel, B. Hauer, *Curr. Opin. Chem. Biol.* **2011**, *15*, 187–193.

[5]   A. Schmid, J. S. Dordick, B. Hauer, A. Kiener, M. Wubbolts, B. Witholt, *Nature* **2001**, *409*, 258–268.

[6]   U. T. Bornscheuer, G. W. Huisman, R. J. Kazlauskas, S. Lutz, J. C. Moore, K. Robins, *Nature* **2012**, *485*, 185–194.

[7]   U. T. Bornscheuer, B. Hauer, K. E. Jaeger, U. Schwaneberg, *Angew. Chemie - Int. Ed.* **2019**, *58*, 36–40.

[8]   T. Davids, M. Schmidt, D. Böttcher, U. T. Bornscheuer, *Curr. Opin. Chem. Biol.* **2013**, *17*, 215–220.

[9]   F. H. Arnold, *Acc. Chem. Res.* **1998**, *31*, 125–131.

[10]   P. A. Romero, F. H. Arnold, *Nat. Rev. Mol. Cell Biol.* **2009**, *10*, 866–876.

[11]   F. H. Arnold, *Angew. Chemie - Int. Ed.* **2019**, *58*, 14420–14426.

[12]   I. V. Korendovych, in *Protein Engineering: Methods and Protocols*, Springer New York, **2018**, pp. 15–23.

[13]   R. Chen, *Trends Biotechnol.* **2001**, *19*, 13–14.

[14]   R. A. Chica, N. Doucet, J. N. Pelletier, *Curr. Opin. Biotechnol.* **2005**, *16*, 378–384.

[15]   S. Lutz, *Curr. Opin. Chem. Biol.* **2010**, *21*, 734–743.

[16]   D. N. Bolon, C. A. Voigt, S. L. Mayo, *Curr. Opin. Chem. Biol.* **2002**, *6*, 125–129.

[17]   H. Kries, R. Blomberg, D. Hilvert, *Curr. Opin. Chem. Biol.* **2013**, *17*, 221–228.

[18]   A. Zanghellini, *Curr. Opin. Biotechnol.* **2014**, *29*, 132–138.

[19]   L. E. Smart, *Alkenes and Aromatics*, Royal Society Of Chemistry, **2002**, 1-178.

[20]   D. Cristina Silva Costa, *Arab. J. Chem.* **2020**, *13*, 799–834.

[21]   T. Katsuki, in *Comprehensive Coordination Chemistry. II*, Elsevier Ltd., **2003**, pp. 207–264.

[22]   C. J. R. Bataille, T. J. Donohoe, *Chem. Soc. Rev.* **2011**, *40*, 114–128.

[23]   D. R. Boyd, G. N. Sheldrake, *Nat. Prod. Rep.* **1998**, *15*, 309–324.

[24]   T. Hudlicky, *ACS Omega* **2018**, *3*, 17326–17340.

[25]   C. Gally, *Enzymatic Asymmetric Dihydroxylation of Alkenes*, University of Stuttgart, Dissertation, **2016**, 1-200.

[26]   J. Halder, *Naphthalen Dioxygenase Aus Pseudomonas Sp. NCIB 9816-4: Systematische Analyse Der Aktiven Tasche*, University of Stuttgart, Dissertation, **2017**, 1-163.

[27]   H. C. Kolb, M. S. VanNieuwenhze, K. B. Sharpless, *Chem. Rev.* **1994**, *94*, 2483–2547.

[28]   K. C. Bhowmick, N. N. Joshi, *Tetrahedron Asymmetry* **2006**, *17*, 1901–1929.

[29]   B. C. Buckland, S. W. Drew, N. C. Connors, M. M. Chartrain, C. Lee, P. M. Salmon, K. Gbewonyo, W. Zhou, P. Gailliot, R. Singhvi, R. C. Olewinski, W. J. Sun, J. Reddy, J. Zhang, B. A. Jackey, C. Taylor, K. E. Goklen, B. Junker, R. L. Greasham, *Metab. Eng.* **1999**, *1*, 63–74.

[30]   T. Hudlicky, D. Gonzalez, D. T. Gibson, *Aldrichimica Acta* **1999**, *32*, 35–62.

[31]   R. Ullrich, M. Hofrichter, *Cell. Mol. Life Sci.* **2007**, *64*, 271–293.

[32]   M. M. Heravi, V. Zadsirjan, M. Esfandyari, T. B. Lashaki, *Tetrahedron Asymmetry* **2017**, *28*, 987–1043.

[33]   B. M. Choudary, N. S. Chowdari, S. Madhi, M. L. Kantam, *J. Org. Chem.* **2003**, *68*, 1736–1746.

[34] M. G. Banwell, M. J. Coster, A. J. Edwards, O. P. Karunaratne, J. A. Smith, L. L. Welling, A. C. Willis, *Aust. J. Chem.* **2003**, *56*, 585–595.

[35] J. S. Reddy, B. V. Rao, *J. Org. Chem.* **2007**, *72*, 2224–2227.

[36] D. R. Boyd, N. D. Sharma, C. J. McGivern, P. J. Stevenson, P. Hoering, C. C. R. Allen, *J. Org. Chem.* **2019**, *84*, 15165–15172.

[37] R. V Ottenbacher, E. P. Talsi, K. P. Bryliakov, *Russ. Chem. Rev.* **2019**, *88*, 1094–1103.

[38] D. Ager, in *Comprehensive Chirality*, Elsevier Ltd., **2012**, pp. 104–128.

[39] M. C. Noe, M. A. Letavic, S. L. Snow, in *Org. React.*, Wiley, **2005**, pp. 109–625.

[40] M. J. Rawling, N. C. O. Tomkinson, *Org. Biomol. Chem.* **2013**, *11*, 1434–1440.

[41] A. B. Zaitsev, H. Adolfsson, *Synthesis (Stuttg).* **2006**, *2006*, 1725–1756.

[42] O. Makowka, *Berichte der Dtsch. Chem. Gesellschaft* **1908**, *41*, 943–944.

[43] K. A. Hofmann, *Berichte der Dtsch. Chem. Gesellschaft* **1912**, *45*, 3329–3336.

[44] R. Criegee, *Liebigs Ann.* **1936**, *522*, 75–96.

[45] R. Criegee, B. Marchand, H. Wonnowius, *Liebigs Ann.* **1942**, *550*, 99–133.

[46] C. Döbler, G. M. Mehltretter, U. Sundermeier, M. Beller, *J. Am. Chem. Soc.* **2000**, *122*, 10289–10297.

[47] V. VanRheenen, R. C. Kelly, D. Y. Cha, *Tetrahedron Lett.* **1976**, *17*, 1973–1976.

[48] M. Schröder, *Chem. Rev.* **1980**, *80*, 187–213.

[49] E. N. Jacobsen, I. Markó, W. S. Mungall, G. Schröder, K. B. Sharpless, *J. Am. Chem. Soc.* **1988**, *110*, 1968–1970.

[50] J. S. M. Wai, I. Marko, J. S. Svendsen, M. G. Finn, E. N. Jacobsen, K. B. Sharpless, *J. Am. Chem. Soc.* **1989**, *111*, 1123–1125.

[51] Y. Ogino, H. Chen, H. L. Kwong, K. B. Sharpless, *Tetrahedron Lett.* **1991**, *32*, 3965–3968.

[52] D. Chang, J. Zhang, B. Witholt, Z. Li, *Biocatal. Biotransformation* **2004**, *22*, 113–131.

[53] L. Wang, K. B. Sharpless, *J. Am. Chem. Soc.* **1992**, *114*, 7568–7570.

[54] S. G. Hentges, K. B. Sharpless, *J. Am. Chem. Soc.* **1980**, *102*, 4263–4265.

[55] K. B. Sharpless, W. Amberg, Y. L. Bennani, G. A. Crispino, J. Hartung, K. S. Jeong, H. L. Kwong, K. Morikawa, Z. M. Wang, D. Xu, X. L. Zhang, *J. Org. Chem.* **1992**, *57*, 2768–2771.

[56] K. Hoi-Lun, C. Sorato, Y. Ogino, C. Hou, K. Barry Sharpless, *Tetrahedron Lett.* **1990**, *31*, 2999–3002.

[57] K. B. Sharpless, *Angew. Chemie - Int. Ed.* **2002**, *41*, 2024–2032.

[58] T. Katsuki, K. B. Sharpless, *J. Am. Chem. Soc.* **1980**, *102*, 5974–5976.

[59] T. K. M. Shing, E. K. W. Tam, V. W. F. Tai, I. H. F. Chung, Q. Jiang, *Chem. - A Eur. J.* **1996**, *2*, 50–57.

[60] T. K. M. Shing, V. W. -F. Tai, E. K. W. Tam, *Angew. Chemie Int. Ed. English* **1994**, *33*, 2312–2313.

[61] B. Plietker, M. Niggemann, *Org. Lett.* **2003**, *5*, 3353–3356.

[62] B. Plietker, M. Niggemann, *J. Org. Chem.* **2005**, *70*, 2402–2405.

[63] B. Plietker, *Synthesis (Stuttg).* **2005**, *15*, 2453–2472.

[64] N. M. Neisius, B. Plietker, *J. Org. Chem.* **2008**, *73*, 3218–3227.

[65] J. Frunzke, C. Loschen, G. Frenking, *J. Am. Chem. Soc.* **2004**, *126*, 3642–3652.

[66] D. E. De Vos, S. De Wildeman, B. F. Sels, P. J. Grobet, P. A. Jacobs, *Angew. Chemie - Int. Ed.* **1999**, *38*, 980–983.

[67] J. W. De Boer, J. Brinksma, W. R. Browne, A. Meetsma, P. L. Alsters, R. Hage, B. L. Feringa, *J. Am. Chem. Soc.* **2005**, *127*, 7990–7991.

[68] T. W. S. Chow, Y. Liu, C. M. Che, *Chem. Commun.* **2011**, *47*, 11204–11206.

[69] S. Enthaler, K. Junge, M. Beller, *Angew. Chemie - Int. Ed.* **2008**, *47*, 3317–3321.

[70] K. Chen, M. Costas, J. Kim, A. K. Tipton, L. Que, *J. Am. Chem. Soc.* **2002**, *124*, 3026–3035.

Literature

[71] M. Borrell, M. Costas, *ACS Sustain. Chem. Eng.* **2018**, *6*, 8410–8416.

[72] Y. Feng, C. Y. Ke, G. Xue, L. Que, *Chem. Commun.* **2009**, 50–52.

[73] C. Zang, Y. Liu, Z. J. Xu, C. W. Tse, X. Guan, J. Wei, J. S. Huang, C. M. Che, *Angew. Chemie - Int. Ed.* **2016**, *55*, 10253–10257.

[74] H. J. H. Fenton, *J. Chem. Soc., Trans.* **1894**, *65*, 899–910.

[75] I. M. Kolthoff, A. I. Medalia, *J. Am. Chem. Soc.* **1949**, *71*, 3777–3783.

[76] R. B. Woodward, F. V. Brutcher, *J. Am. Chem. Soc.* **1958**, *80*, 209–211.

[77] L. Emmanuvel, T. M. Ali Shaikh, A. Sudalai, *Org. Lett.* **2005**, *7*, 5071–5074.

[78] W. Zhong, J. Yang, X. Meng, Z. Li, *J. Org. Chem.* **2011**, *76*, 9997–10004.

[79] M. Fujita, M. Wakita, T. Sugimura, *Chem. Commun.* **2011**, *47*, 3983–3985.

[80] V. L. Paddock, R. J. Phipps, A. Conde-Angulo, A. Blanco-Martin, C. Giró-Mañas, L. J. Martin, A. J. P. White, A. C. Spivey, *J. Org. Chem.* **2011**, *76*, 1483–1486.

[81] T. D. Bugg, in *Encyclopedia of Life Sciences*, Wiley, **2001**, pp. 1–8.

[82] F. F. Özgen, S. Schmidt, *Biocatalysis* **2019**, 57–82.

[83] S. Resnick, K. Lee, D. Gibson, *J. Ind. Microbiol. Biotechnol.* **1996**, *17*, 438–457.

[84] M. A. Endoma, V. P. Bui, J. Hansen, T. Hudlicky, *Org. Process Res. Dev.* **2002**, *6*, 525–532.

[85] F. Fabris, J. Collins, B. Sullivan, H. Leisch, T. Hudlicky, *Org. Biomol. Chem.* **2009**, *7*, 2619–2627.

[86] M. G. Quintana, H. Dalton, *Enzyme Microb. Technol.* **1999**, *24*, 232–236.

[87] O. Kweon, S. J. Kim, S. Baek, J. C. Chae, M. D. Adjei, D. H. Baek, Y. C. Kim, C. E. Cerniglia, *BMC Biochem.* **2008**, *9*, 11.

[88] D. T. Gibson, J. R. Koch, R. E. Kallio, *Biochemistry* **1968**, *7*, 2653–2662.

[89] D. T. Gibson, J. R. Koch, C. L. Schuld, R. E. Kallio, *Biochemistry* **1968**, *7*, 3795–3802.

[90] H. J. Heipieper, P. M. Martínez, in *Handbook of Hydrocarbon and Lipid Microbiology*, Springer Berlin Heidelberg, **2010**, pp. 1563–1573.

[91] P. C. K. Lau, H. Bergeron, D. Labbé, Y. Wang, R. Brousseau, D. T. Gibson, *Gene* **1994**, *146*, 7–13.

[92] D. T. Gibson, M. Hensley, H. Yoshioka, T. J. Mabry, *Biochemistry* **1970**, *9*, 1626–1630.

[93] D. T. Gibson, G. E. Cardini, F. C. Maseles, R. E. Kallio, *Biochemistry* **1970**, *9*, 1631–1635.

[94] W. K. Yeh, D. T. Gibson, T.-N. Liu, *Biochem. Biophys. Res. Commun.* **1977**, *78*, 401–410.

[95] G. J. Zylstra, W. R. McCombie, D. T. Gibson, B. A. Finette, *Appl. Environ. Microbiol.* **1988**, *54*, 1498–1503.

[96] Y. Wang, M. Rawlings, D. T. Gibson, D. Labbé, H. Bergeron, R. Brousseau, P. C. K. Lau, *MGG Mol. Gen. Genet.* **1995**, *246*, 570–579.

[97] R. E. Parales, J. V. Parales, D. A. Pelletier, J. L. Ditty, in *Advances in Applied Microbiology*, Elsevier Ltd., **2008**, pp. 1–73.

[98] R. E. Parales, J. L. Ditty, C. S. Harwood, *Appl. Environ. Microbiol.* **2000**, *66*, 4098–4104.

[99] G. J. Zylstra, D. T. Gibson, *J. Biol. Chem.* **1989**, *264*, 14940–14946.

[100] C. Batie, D. Ballou, C. Corell, in *Chemistry and Biochemistry of Flavoenzymes Volume III*, CRC Press, **1991**, pp. 543–556.

[101] J. Chakraborty, D. Ghosal, A. Dutta, T. K. Dutta, *J. Biomol. Struct. Dyn.* **2012**, *30*, 419–436.

[102] R. Friemann, K. Lee, E. N. Brown, D. T. Gibson, H. Eklund, S. Ramaswamy, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2009**, *65*, 24–33.

[103] D. J. Ferraro, L. Gakhar, S. Ramaswamy, *Biochem. Biophys. Res. Commun.* **2005**, *338*, 175–190.

[104] Y. Hurtubise, D. Barriault, M. Sylvestre, *J. Bacteriol.* **1998**, *180*, 5828–5835.

[105] C. L. Colbert, M. M. J. Couture, L. D. Eltis, J. T. Bolin, *Structure* **2000**, *8*, 1267–1278.

[106] D. Schneider, C. L. Schmidt, *Biochim. Biophys. Acta - Bioenerg.* **2005**, *1710*, 1–12.

[107] B. Kauppi, K. Lee, E. Carredano, R. E. Parales, D. T. Gibson, H. Eklund, S. Ramaswamy, *Structure* **1998**, *6*, 571–586.

[108] T. Ohta, S. Chakrabarty, J. D. Lipscomb, E. I. Solomon, *J. Am. Chem. Soc.* **2008**, *130*, 1601–1610.

[109] S. M. Barry, G. L. Challis, *ACS Catal.* **2013**, *3*, 2362–2370.

[110] Y. Ashikawa, Z. Fujimoto, Y. Usami, K. Inoue, H. Noguchi, H. Yamane, H. Nojiri, *BMC Struct. Biol.* **2012**, *12*, 15.

[111] T. D. H. Bugg, *Tetrahedron* **2003**, *59*, 7075–7101.

[112] B. S. Rivard, M. S. Rogers, D. J. Marell, M. B. Neibergall, S. Chakrabarty, C. J. Cramer, J. D. Lipscomb, *Biochemistry* **2015**, *54*, 4652–4664.

[113] A. Karlsson, J. V. Parales, R. E. Parales, D. T. Gibson, H. Eklund, S. Ramaswamy, *Science* **2003**, *299*, 1039–1042.

[114] M. D. Wolfe, J. V. Parales, D. T. Gibson, J. D. Lipscomb, *J. Biol. Chem.* **2001**, *276*, 1945–1953.

[115] M. Costas, M. P. Mehn, M. P. Jensen, L. Que, *Chem. Rev.* **2004**, *104*, 939–986.

[116] M. M. Abu-Omar, A. Loaiza, N. Hontzeas, *Chem. Rev.* **2005**, *105*, 2227–2252.

[117] S. Chakrabarty, R. N. Austin, D. Deng, J. T. Groves, J. D. Lipscomb, *J. Am. Chem. Soc.* **2007**, *129*, 3514–3515.

[118] K. D. Sutherlin, B. S. Rivard, L. H. Böttger, L. V. Liu, M. S. Rogers, M. Srnec, K. Park, Y. Yoda, S. Kitao, Y. Kobayashi, M. Saito, M. Seto, M. Hu, J. Zhao, J. D. Lipscomb, E. I. Solomon, *J. Am. Chem. Soc.* **2018**, *140*, 5544–5559.

[119] C. Perry, E. L. C. De Los Santos, L. M. Alkhalaf, G. L. Challis, *Nat. Prod. Rep.* **2018**, *35*, 622–632.

[120] J. Joern, T. Sakamoto, A. Arisawa, F. Arnold, *J. Biomol. Screen.* **2001**, *6*, 219–223.

[121] D. R. Boyd, N. D. Sharma, P. J. Stevenson, J. Chima, D. J. Gray, H. Dalton, *Tetrahedron Lett.* **1991**, *32*, 3887–3890.

[122] S. M. Resnick, D. T. Gibson, *Biodegradation* **1993**, *4*, 195–203.

[123] D. R. Boyd, D. S. Sharma, N. A. Kerley, R. A. McMoride, G. N. Sheldrake, P. Williams, H. Dalton, *J. Chem. Soc.* **1996**, 67–74.

[124] D. R. Boyd, N. D. Sharma, S. A. Haughey, M. A. Kennedy, B. T. McMurray, G. N. Sheldrake, C. C. R. Allen, H. Dalton, K. Sproule, *J. Chem. Soc. Perkin Trans. 1* **1998**, 1929–1934.

[125] M. A. Vila, V. Steck, S. Rodriguez Giordano, I. Carrera, R. Fasan, *ChemBioChem* **2020**, *21*, 1981–1987.

[126] P. Höring, K. Rothschild-Mancinelli, N. D. Sharma, D. R. Boyd, C. C. R. Allen, *J. Mol. Catal. B Enzym.* **2016**, *134*, 396–406.

[127] C. C. Lange, L. P. Wackett, *J. Bacteriol.* **1997**, *179*, 3858–3865.

[128] T. Hudlicky, R. Fan, H. Luna, H. Olivo, J. Price, *Pure Appl. Chem.* **1992**, *64*, 1109–1113.

[129] L. P. Wackett, *Enzyme Microb. Technol.* **2002**, *31*, 577–587.

[130] D. R. Boyd, N. D. Sharma, C. C. R. Allen, *Curr. Opin. Biotechnol.* **2001**, *12*, 564–573.

[131] T. Hudlicky, H. F. Olivo, *J. Am. Chem. Soc.* **1992**, *114*, 9694–9696.

[132] M. A. A. Endoma-Arias, J. R. Hudlicky, R. Simionescu, T. Hudlicky, *Adv. Synth. Catal.* **2014**, *356*, 333–339.

[133] F. Tang, P. Lan, B. Bolte, M. G. Banwell, J. S. Ward, A. C. Willis, *J. Org. Chem.* **2018**, *83*, 14049–14056.

[134] A. Dror, A. Fishman, *Comput. Struct. Biotechnol. J.* **2012**, *2*, e201209011.

[135] F. Catteral, P. Williams, *J. Gen. Microbiol.* **1971**, *67*, 117–124.

[136] S. V. Ley, F. Sternfeld, S. Taylor, *Tetrahedron Lett.* **1987**, *28*, 225–226.

Literature

[137] D. G. H. Ballard, A. Courtis, I. M. Shirley, S. C. Taylor, *J. Chem. Soc. Chem. Commun.* **1983**, 954–955.

[138] S. C. Taylor, *Biochemical Process*, **1995**, Patent US4508822A.

[139] M. A. Vila, M. Brovetto, D. Gamenara, P. Bracco, G. Zinola, G. Seoane, S. Rodríguez, I. Carrera, *J. Mol. Catal. B Enzym.* **2013**, *96*, 14–20.

[140] J. J. Shie, J. M. Fang, C. H. Wong, *Angew. Chemie - Int. Ed.* **2008**, *47*, 5788–5791.

[141] T. Hudlicky, G. Seoane, T. Pettus, *J. Org. Chem.* **1989**, *54*, 4239–4243.

[142] B. Ensley, B. Ratzkin, T. Osslund, M. Simon, L. Wackett, D. Gibson, *Science* **1983**, *222*, 167–169.

[143] W. Weyler, T. C. Dodge, J. J. Lauff, D. J. Wendt, *Microbiol Production of Indigo*, **1995**, Patent US6190892B1.

[144] J. Y. Kim, K. Lee, Y. Kim, C. K. Kim, K. Lee, *Lett. Appl. Microbiol.* **2003**, *36*, 343–348.

[145] J. L. Wissner, W. Escobedo-Hinojosa, P. M. Heinemann, A. Hunold, B. Hauer, in *Methods Enzymol.*, **2020**, pp. 63–93.

[146] H. Jiang, R. E. Parales, N. A. Lynch, D. T. Gibson, *J. Bacteriol.* **1996**, *178*, 3133–3139.

[147] S. Beil, J. R. Mason, K. N. Timmis, D. H. Pieper, *J. Bacteriol.* **1998**, *180*, 5520–5528.

[148] R. E. Parales, S. M. Resnick, C. L. Yu, D. R. Boyd, N. D. Sharma, D. T. Gibson, *J. Bacteriol.* **2000**, *182*, 5495–5504.

[149] N. Zhang, B. G. Stewart, J. C. Moore, R. L. Greasham, D. K. Robinson, B. C. Buckland, C. Lee, *Metab. Eng.* **2000**, *2*, 339–348.

[150] T. Sakamoto, J. M. Joern, A. Arisawa, F. H. Arnold, *Appl. Environ. Microbiol.* **2001**, *67*, 3882–3887.

[151] L. M. Newman, H. Garcia, T. Hudlicky, S. A. Selifonov, *Tetrahedron* **2004**, *60*, 729–734.

[152] C. Gally, B. M. Nestl, B. Hauer, *Angew. Chemie - Int. Ed.* **2015**, *54*, 12952–12956.

[153] M. A. Vila, D. Umpiérrez, N. Veiga, G. Seoane, I. Carrera, S. Rodríguez Giordano, *Adv. Synth. Catal.* **2017**, *359*, 2149–2157.

[154] J. M. Halder, B. M. Nestl, B. Hauer, *ChemCatChem* **2018**, *10*, 178–182.

[155] P. M. Heinemann, D. Armbruster, B. Hauer, *Nat. Commun.* **2021**, *12*, 1–12.

[156] J. L. Wissner, J. Ludwig, W. Escobedo-Hinojosa, B. Hauer, *J. Biotechnol.* **2020**, *325*, 380–388.

[157] J. L. Wissner, W. Escobedo-Hinojosa, A. Vogel, B. Hauer, *J. Biotechnol.* **2021**, *326*, 37–39.

[158] J. L. Wissner, J. T. Schelle, W. Escobedo-Hinojosa, A. Vogel, B. Hauer, *Adv. Synth. Catal.* **2021**, *363*, DOI 10.1002/adsc.202100296.

[159] D. R. Boyd, N. D. Sharma, P. L. Loke, J. G. Carroll, P. J. Stevenson, P. Hoering, C. C. R. Allen, *Front. Bioeng. Biotechnol.* **2021**, *8*, 1568–1584.

[160] D. Boyd, N. Sharma, I. Brannigan, C. McGivern, P. Nockemann, P. Stevenson, C. McRoberts, P. Hoering, C. Allen, *Adv. Synth. Catal.* **2019**, *361*, 2526–2537.

[161] M. A. Vila, D. Umpiérrez, G. Seoane, S. Rodríguez, I. Carrera, N. Veiga, *J. Mol. Catal. B Enzym.* **2016**, *133*, S410–S419.

[162] G. L. Rosano, E. A. Ceccarelli, *Front. Microbiol.* **2014**, *5*, 172–189.

[163] V. P. Bui, T. Vidar Hansen, Y. Stenstrøm, T. Hudlicky, D. W. Ribbons, *New J. Chem.* **2001**, *25*, 116–124.

[164] J. F. Trant, J. Froese, T. Hudlicky, *Tetrahedron Asymmetry* **2013**, *24*, 184–190.

[165] J. Shiloach, J. Kaufman, A. S. Guillard, R. Fass, *Biotechnol. Bioeng.* **1996**, *49*, 421–428.

[166] S.-P. Ouyang, S.-Y. Sun, Q. Liu, J. Chen, G.-Q. Chen, *Appl. Microbiol. Biotechnol.* **2007**, *74*, 43–49.

[167] C. C. R. Allen, D. R. Boyd, H. Dalton, N. D. Sharma, S. A. Haughey, R. A. S. McMordie, B. T. McMurray, G. N. Sheldrake, K. Sproule, *J. Chem. Soc. Chem. Commun.* **1995**, 119–120.

[168] L.-M. Guzman, D. Belin, M. Carson, J. Beckwith, *J. Bacteriol.* **1995**, *177*, 4121–4130.

[169] D. G. Gibson, *Methods Enzymol.* **2011**, *498*, 349–361.

[170] K. A. Datsenko, B. L. Wanner, *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 6640–6645.

[171] T. Baba, T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, M. Baba, K. A. Datsenko, M. Tomita, B. L. Wanner, H. Mori, *Mol. Syst. Biol.* **2006**, *2*, 2006.0008.

[172] C. Bagnéris, R. Cammack, J. R. Mason, C. Bagne, *Appl. Environ. Microbiol.* **2005**, *71*, 1570–1580.

[173] A. Masumoto, T. Kikuchi, N. Ono, H. Uno, H. Nakashima, *Compound and Method of Producing Organic Semiconductor Device*, **2011**, Patent US7928221B2.

[174] J. D. Ludwig, *Studien Zur TDO Katalysierten Dihydroxylierung von Aromaten*, University of Stuttgart, Master Thesis, **2019**, 1-91.

[175] R. Agarwala, T. Barrett, J. Beck, D. A. Benson, C. Bollin, E. Bolton, D. Bourexis, J. R. Brister, S. H. Bryant, K. Canese, M. Cavanaugh, C. Charowhas, K. Clark, I. Dondoshansky, M. Feolo, L. Fitzpatrick, K. Funk, L. Y. Geer, V. Gorelenkov, A. Graeff, W. Hlavina, B. Holmes, M. Johnson, B. Kattman, V. Khotomlianski, A. Kimchi, M. Kimelman, M. Kimura, P. Kitts, W. Klimke, A. Kotliarov, S. Krasnov, A. Kuznetsov, M. J. Landrum, D. Landsman, S. Lathrop, J. M. Lee, C. Leubsdorf, Z. Lu, T. L. Madden, A. Marchler-Bauer, A. Malheiro, P. Meric, I. Karsch-Mizrachi, A. Mnev, T. Murphy, R. Orris, J. Ostell, C. O'Sullivan, V. Palanigobu, A. R. Panchenko, L. Phan, B. Pierov, K. D. Pruitt, K. Rodarmer, E. W. Sayers, V. Schneider, C. L. Schoch, G. D. Schuler, S. T. Sherry, K. Siyan, A. Soboleva, V. Soussov, G. Starchenko, T. A. Tatusova, F. Thibaud-Nissen, K. Todorov, B. W. Trawick, D. Vakatov, M. Ward, E. Yaschenko, A. Zasypkin, K. Zbicz, *Nucleic Acids Res.* **2018**, *46*, D8–D13.

[176] E. Díaz, A. Ferrández, J. L. García, *J. Bacteriol.* **1998**, *180*, 2915–2923.

[177] M. Sakaitani, F. Rusnak, N. R. Quinn, C. Tu, T. B. Frigo, G. A. Berchtold, C. T. Walsh, *Biochemistry* **1990**, *29*, 6789–6798.

[178] I. M. Keseler, A. Mackie, A. Santos-Zavaleta, R. Billington, C. Bonavides-Martínez, R. Caspi, C. Fulcher, S. Gama-Castro, A. Kothari, M. Krummenacker, M. Latendresse, L. Muñiz-Rascado, Q. Ong, S. Paley, M. Peralta-Gil, P. Subhraveti, D. A. Velázquez-Ramírez, D. Weaver, J. Collado-Vides, I. Paulsen, P. D. Karp, *Nucleic Acids Res.* **2017**, *45*, D543–D550.

[179] J. L. Wissner, J. Ludwig, W. Escobedo-Hinojosa, B. Hauer, *MethodsX* **2020**, *7*, 101143.

[180] K. P. Y. Fong, H. M. Tan, *FEBS Lett.* **1999**, *451*, 5–9.

[181] J. J. Kelley, E. E. Dekker, *J. Biol. Chem.* **1984**, *259*, 2124–2129.

[182] A. Bateman, M. J. Martin, S. Orchard, M. Magrane, R. Agivetova, S. Ahmad, E. Alpi, E. H. Bowler-Barnett, R. Britto, B. Bursteinas, H. Bye-A-Jee, R. Coetzee, A. Cukura, A. Da Silva, P. Denny, T. Dogan, T. G. Ebenezer, J. Fan, L. G. Castro, P. Garmiri, G. Georghiou, L. Gonzales, E. Hatton-Ellis, A. Hussein, A. Ignatchenko, G. Insana, R. Ishtiaq, P. Jokinen, V. Joshi, D. Jyothi, A. Lock, R. Lopez, A. Luciani, J. Luo, Y. Lussi, A. MacDougall, F. Madeira, M. Mahmoudy, M. Menchi, A. Mishra, K. Moulang, A. Nightingale, C. S. Oliveira, S. Pundir, G. Qi, S. Raj, D. Rice, M. R. Lopez, R. Saidi, J. Sampson, T. Sawford, E. Speretta, E. Turner, N. Tyagi, P. Vasudev, V. Volynkin, K. Warner, X. Watkins, R. Zaru, H. Zellner, A. Bridge, S. Poux, N. Redaschi, L. Aimo, G. Argoud-Puy, A. Auchincloss, K. Axelsen, P. Bansal, D. Baratin, M. C. Blatter, J. Bolleman, E. Boutet, L. Breuza, C. Casals-Casas, E. de Castro, K. C. Echioukh, E. Coudert, B. Cuche, M. Doche, D. Dornevil, A. Estreicher, M. L. Famiglietti, M. Feuermann, E. Gasteiger, S. Gehant, V. Gerritsen, A. Gos, N. Gruaz-Gumowski, U. Hinz, C. Hulo, N. Hyka-Nouspikel, F. Jungo, G. Keller, A. Kerhornou, V. Lara, P. Le Mercier, D. Lieberherr, T. Lombardot, X. Martin, P. Masson, A. Morgat, T. B. Neto, S. Paesano, I. Pedruzzi, S. Pilbout, L. Pourcel, M. Pozzato, M. Pruess, C. Rivoire, C. Sigrist, K. Sonesson, A. Stutz, S. Sundaram, M. Tognolli, L. Verbregue, C. H. Wu, C. N. Arighi,

L. Arminski, C. Chen, Y. Chen, J. S. Garavelli, H. Huang, K. Laiho, P. McGarvey, D. A. Natale, K. Ross, C. R. Vinayaka, Q. Wang, Y. Wang, L. S. Yeh, J. Zhang, *Nucleic Acids Res.* **2021**, *49*, D480–D489.

[183] K. Shindo, R. Nakamura, A. Osawa, O. Kagami, K. Kanoh, K. Furukawa, N. Misawa, *J. Mol. Catal. B Enzym.* **2005**, *35*, 134–141.

[184] M. Kwit, J. Gawronski, D. R. Boyd, N. D. Sharma, M. Kaik, R. A. More O'Ferrall, J. S. Kudavalli, *Chem. - A Eur. J.* **2008**, *14*, 11500–11511.

[185] D. M. Jerina, J. W. Daly, A. M. Jeffrey, D. T. Gibson, *Arch. Biochem. Biophys.* **1971**, *142*, 394–396.

[186] F. Orsini, G. Sello, E. Travaini, P. Di Gennaro, *Tetrahedron Asymmetry* **2002**, *13*, 253–259.

[187] F. Orsini, G. Sello, S. Bernasconi, G. Fallacara, *Tetrahedron Lett.* **2004**, *45*, 9253–9255.

[188] A. M. Jeffrey, H. J. C. Yeh, D. M. Jerina, *J. Org. Chem.* **1974**, *39*, 1405–1407.

[189] P. Mukherjee, S. J. S. Roy, T. K. Sarkar, *Org. Lett.* **2010**, *12*, 2472–2475.

[190] A. M. McIver, S. V. B. J. Garikipati, K. S. Bankole, M. Gyamerah, T. L. Peeples, *Biotechnol. Prog.* **2008**, *24*, 593–598.

[191] A. M. Jeffrey, H. J. C. Yeh, D. M. Jerina, T. R. Patel, J. F. Davey, D. T. Gibson, *Biochemistry* **1975**, *14*, 575–584.

[192] C. C. R. Allen, D. R. Boyd, H. Dalton, N. D. Sharma, I. Brannigan, N. A. Kerley, G. N. Sheldrake, S. C. Taylor, *J. Chem. Soc. Chem. Commun.* **1995**, 117–118.

[193] R. E. Parales, D. Gibson, S. Resnick, K. Lee, *Novel Naphthalene Dioxygenase And Methods For Their Use*, **2000**, Patent WO2000037480A1.

[194] D. R. Boyd, N. D. Sharma, L. V Modyanova, J. G. Carroll, J. F. Malone, C. C. Allen, J. T. Hamilton, D. T. Gibson, R. E. Parales, H. Dalton, *Can. J. Chem.* **2002**, *80*, 589–600.

[195] D. R. Boyd, N. D. Sharma, G. P. Coen, F. Hempenstall, V. Ljubez, J. F. Malone, C. C. R. Allen, J. T. G. Hamilton, *Org. Biomol. Chem.* **2008**, *6*, 3957–3966.

[196] E. Chovancova, A. Pavelka, P. Benes, O. Strnad, J. Brezovsky, B. Kozlikova, A. Gora, V. Sustr, M. Klvana, P. Medek, L. Biedermannova, J. Sochor, J. Damborsky, *PLoS Comput. Biol.* **2012**, *8*, 23–30.

[197] M. Mohammadi, J. F. Viger, P. Kumar, D. Barriault, J. T. Bolin, M. Sylvestre, *J. Biol. Chem.* **2011**, *286*, 27612–27621.

[198] J. Schelle, *Enzym-Engineering der Toluen-Dioxygenase aus Pseudomonas putida hinsichtlich N-heterozyklischer Aromaten*, University of Stuttgart, Master Thesis, **2020** 1-92.

[199] D. S. Torok, S. M. Resnick, J. M. Brand, D. L. Cruden, D. T. Gibson, *J. Bacteriol.* **1995**, *177*, 5799–5805.

[200] D. R. Boyd, N. D. Sharma, M. R. J. Dorrity, M. V. Hand, M. R. A. S., J. F. Malone, H. P. Porter, H. Dalton, J. Chima, G. N. Sheldrake, **1993**, *J. Chem. S*, 1065–1071.

[201] V. Sridharan, P. A. Suryavanshi, J. C. Menéndez, *Chem. Rev.* **2011**, *111*, 7157–7259.

[202] M. Nettekoven, J. Fingerle, U. Grether, S. Grüner, A. Kimbara, B. Püllmann, M. Rogers-Evans, S. Röver, F. Schuler, T. Schulz-Gasch, C. Ullmer, *Bioorganic Med. Chem. Lett.* **2013**, *23*, 1177–1181.

[203] D. Zheng, M. Yang, J. Zhuo, K. Li, H. Zhang, J. Yang, B. Cui, Y. Chen, *J. Mol. Catal. B Enzym.* **2014**, *110*, 87–91.

[204] C. Yin, X. Q. Dong, X. Zhang, *Adv. Synth. Catal.* **2018**, *360*, 4319–4324.

[205] N. V. Suryanarayana Birudukota, R. Franke, B. Hofer, *Org. Biomol. Chem.* **2016**, *14*, 3821–3837.

[206] M. D. Garrett, R. Scott, G. N. Sheldrake, H. Dalton, P. Goode, *Org. Biomol. Chem.* **2006**, *4*, 2710.

[207] E. L. Ang, J. P. Obbard, H. Zhao, *FEBS J.* **2007**, *274*, 928–939.

[208] R. E. Parales, K. Lee, S. M. Resnick, H. Jiang, D. J. Lessner, D. T. Gibson, *J. Bacteriol.*

2000, *182*, 1641–1649.

[209] C. Preston-Herrera, A. S. Jackson, B. O. Bachmann, J. T. Froese, *Org. Biomol. Chem.* **2021**, *19*, 775–784.

[210] J. Li, J. Min, Y. Wang, W. Chen, Y. Kong, T. Guo, J. K. Mahto, M. Sylvestre, X. Hu, *Appl. Environ. Microbiol.* **2020**, *86*, 1–17.

[211] M. T. Reetz, M. Bocola, J. D. Carballeira, D. Zha, A. Vogel, *Angew. Chemie - Int. Ed.* **2005**, *44*, 4192–4196.

[212] S. Kille, C. G. Acevedo-Rocha, L. P. Parra, Z. G. Zhang, D. J. Opperman, M. T. Reetz, J. P. Acevedo, *ACS Synth. Biol.* **2013**, *2*, 83–92.

[213] G. Sello, F. Orsini, *Mini. Rev. Org. Chem.* **2005**, *1*, 77–92.

[214] F. Feyza Özgen, M. E. Runda, B. O. Burek, P. Wied, J. Z. Bloh, R. Kourist, S. Schmidt, *Angew. Chemie - Int. Ed.* **2020**, *59*, 3982–3987.

[215] J. Alonso-Gutierrez, R. Chan, T. S. Batth, P. D. Adams, J. D. Keasling, C. J. Petzold, T. S. Lee, *Metab. Eng.* **2013**, *19*, 33–41.

[216] N. Oberleitner, A. K. Ressmann, K. Bica, P. Gärtner, M. W. Fraaije, U. T. Bornscheuer, F. Rudroff, M. D. Mihovilovic, *Green Chem.* **2017**, *19*, 367–371.

6 RESEARCH ARTICLES

Hereinafter, the original unprocessed research articles I, II and III (corresponding to publications 2, 4 and 5) are listed, together with the appertaining supplementary information.

6.1 RESEARCH ARTICLE I

Manuscript

An enhanced toluene dioxygenase platform for the production of *cis*-1,2-dihydrocatechol in *Escherichia coli* BW25113 lacking glycerol dehydrogenase activity

Julian L. Wissner, Julian Ludwig, Wendy Escobedo-Hinojosa, Bernhard Hauer *

Institute of Technical Biochemistry, University of Stuttgart, Allmandring 31, 70569 Stuttgart, Germany

ABSTRACT

The compound *cis*-1,2-dihydrocatechol (DHC) is highly valuable since it finds wide application in the production of fine chemicals and bioactive compounds with medical relevance. The biotechnological process to generate DHC involves a dearomatizing dihydroxylation reaction catalyzed by toluene dioxygenase (TDO) from *P. putida* F1, employing benzene as substrate. We aimed to enhance the biotechnological *E. coli* BW25113 platform for DHC production by identifying the key operational parameters positively influencing the final isolated yield. Thereby, we observed an unreported downstream reaction, generating catechol from DHC, affecting, in a negative manner, the final titer for the product. Expression temperature for the TDO-system showed to have the highest influence in terms of final isolated yield. A KEIO-collection-based screening approach highlighted glycerol dehydrogenase (GldA) as the main responsible enzyme for the undesired reaction. We transferred the TDO-system to *E. coli* BW25113 Δ*gldA* and applied the enhanced operational set-up on it. This enhanced platform enabled the production of 1.41 g L^{-1} DHC in isolated yield, which represents a two-fold increase compared with the starting working conditions. To our knowledge, this is the highest DHC production accomplished in recombinant *E. coli* at semi-preparative scale, providing a robust and accessible biotechnological platform for DHC synthesis.

1. Introduction

The compound *cis*-1,2-dihydrocatechol (DHC), also referred as *cis*-3,5-cyclohexadien-1,2-diol is highly valuable. The DHC scaffold finds extensive application in the production of a variety of fine chemicals and bioactive compounds. For instance, this notable *cis*-dihydrodiol is employed for the production of polyphenylene (Fig. 1A; **E1**), (Ballard et al., 1983; Taylor, 1995), and for the synthesis of precursors aimed to produce organic semiconductors (Fig. 1A; **E2**), (Masumoto et al., 2011). In the pharmaceutical field DHC is employed as a backbone to generate relevant pharmaceuticals such as, (+)-lycoricidine for cancer treatment (Fig. 1A; **E3**), (Hudlicky and Olivo, 1992), and the anti-diabetic agent (+)-pinitol (Fig. 1A; **E4**), (Ley et al., 1987). The first organic synthesis of DHC was established by Nakajima in 1959. Such process involves a potassium permanganate driven oxidation of 3,4,5,6-tetrachlorocyclohexene, followed by a zinc-powder catalyzed dechlorination step, enabling an isolated yield of 41 % (Nakajima et al., 1959). Later on, Paddock and colleagues, developed a more efficient synthesis route to

generate acetonide-protected *meso*-1,2-dihydrocatechol. They employed 1,2-cyclohexadiene as substrate in a three step process, giving a 50 % overall yield (Paddock et al., 2011). Some disadvantages regarding the organic synthesis of DHC are in addition to the need of expensive substrates, the use of toxic heavy metal catalysts and organic solvents, long reaction times, as well as, reduced isolated yields.

In general, the conversion of aromatic compounds to nonaromatic molecules is a challenging reaction, due to the strong stability of the aromatic ring-system (Quintana and Dalton, 1999). Thus, it is extraordinary, how in nature it is possible to find enzymes such as Rieske non-heme iron dioxygenases (ROs) capable to overcome such chemical stability. ROs are capable of dihydroxylating aromatics by using molecular oxygen as co-substrate. Thus, enabling the access to chiral *cis*-dihydrodiols, not only in high yields and outstanding enantiomeric excess, but also under environmentally friendly conditions compared to the heavy metal dependent chemical routes. Indeed, there is not a single one-step chemical reaction yet, which grants the direct dihydroxylation of benzene to this valuable chiral *cis*-diol. One of the most widely studied

---

* Corresponding author.
  E-mail address: bernhard.hauer@itb.uni-stuttgart.de (B. Hauer).

RO is toluene dioxygenase (TDO) from *Pseudomonas putida* F1. The broad substrate scope for TDO comprises over 100 non-natural compounds ranging from aromatics, including benzene, heteroaromatics, poly- and heterocyclic arenes, to halogenated alkenes (Parales and Resnick, 2006). There are several reports on TDO, covering important aspects regarding the system characterization. Nevertheless, there are only a few works exploiting the capability of TDO to perform the direct dearomatizing dihydroxylation of benzene, generating DHC in one single-step reaction. Such potential should be foster in order to produce important amounts of this demanding product.

Initially, the pioneering work of Gibson and colleagues in 1968, led to the identification of the intermediate DHC in the benzene degradation pathway in *P. putida* F1 (Gibson et al., 1968). The genes encoding the upper degradation pathway in this microorganism form part of the *tod* operon, whereby the acronym *tod* stands for "toluene dioxygenase", since tolune is the natural substrate (Zylstra et al., 1988). The multicomponent TDO system was elucidated as the responsible for such reaction (Yeh et al., 1977). This system is comprising a flavoprotein reductase (TodA), a ferredoxin (TodB) and a terminal dioxygenase consisting out of two subunits TodC1 and TodC2 (α- and β- subunits, respectively). In addition, they successfully achieved the heterologous expression of the *tod* genes encoding for the TDO multicomponent system using *E. coli* JM109 (pDTG601) as host strain (Zylstra and Gibson, 1989). There are a few reports regarding the development of biotechnological approaches for the biotransformation of benzene to the corresponding *cis*-diol, employing *P. putida* strains or recombinant systems in *E. coli* and *Aeromonas hydrophilia* strains as biosynthetic platform. Table 1 condenses the research journey, reporting the production of DHC by using TDO as biocatalyst, focused on the bacterial strain, working set-up, and achieved yield.

Despite the actual potential of TDO as biocatalyst for the synthesis of DHC, biotechnological efforts to enhance the production in recombinant platforms stopped more than one decade ago. In order to further exploit this promising approach, we aimed to improve the TDO biotechnological platform in the *E. coli* BW25113 strain, for the generation of DHC. This work provides a sustainable, low-cost, and environmentally friendly process for the biosynthesis of this chemically attractive scaffold.

2. Material and methods

2.1. Materials

Chemicals and solvents used in this work were obtained at the highest purity degree available from Sigma-Aldrich (St. Louis, US) and Carl Roth (Karlsruhe, DE). Since DHC is not commercially available, it was biosynthesized according to the procedure described in section 2.9 below.

2.2. Bacterial strains and growth conditions

E. coli strains and their corresponding genotype are provided in Table 2. The strain XL-1 Blue was employed for cloning purposes, while strain BW25113 was used as expression host. Strain JM109 (DE3) was exclusively employed for comparative purposes during the examination of DHC dehydrogenation in *E. coli*. All *E. coli* strains were routinely grown in LB medium at 37 °C and 180 rpm. When required, ampicillin was added at a final concentration of $100\,mg\,L^{-1}$. In addition, we also employed 88 different dehydrogenase knock-out strains from the KEIO collection (Table S1), to perform a screening approach to detect DHC consumption.

2.3. pBAD18-TDO plasmid construction

Plasmid details and primer sequences are provided in Tables 3 and 4, respectively. The expression vector pBAD18-TDO was constructed as follows; pBAD18 plasmid was employed as backbone to harbor the toluene dioxygenase multicomponent system, from *P. putida* F1, encoded in the expression vector pDTG601 (Zylstra and Gibson, 1989). KOD HS polymerase was employed to perform all PCRs following the protocol described by the supplier (Novagen, Madison, US). The pBAD18 backbone was amplified by using primers 1 and 2 *via* PCR. The insert containing the whole TDO multicomponent system was amplified by overlap extension PCR with the primer set 3−10. Backbone and insert were ligated following the Gibson assembly protocol (Gibson, 2011). The resulting plasmid was transformed into chemical competent *E. coli* XL-1 Blue cells (Hanahan, 1983). Afterwards, plasmid was extracted with the *Zyppy Plasmid Miniprep Kit* (Zymo research, Tustin, US), and submitted for Sanger sequencing (GATC-Biotech, Konstanz, DE), to assess in frame positioning of the insert and correct amplification in terms of sequence and length.

2.4. Expression conditions for the pBAD18-TDO system

For expression purposes, LB overnight precultures were prepared inoculating a single colony of freshly transformed *E. coli* BW25113 cells with the vector system pBAD18-TDO. Protein expression was performed in baffled Erlenmeyer flasks (2 L), containing 500 mL TB medium supplemented with ampicillin. Flasks were inoculated with the overnight precultures to an OD_{600} of 0.05 and incubated at 37 °C and at 100 rpm. When an optical density OD_{600} of 0.7−0.9 was reached, a final concentration of 10 mM of L-arabinose was added in order to induce protein
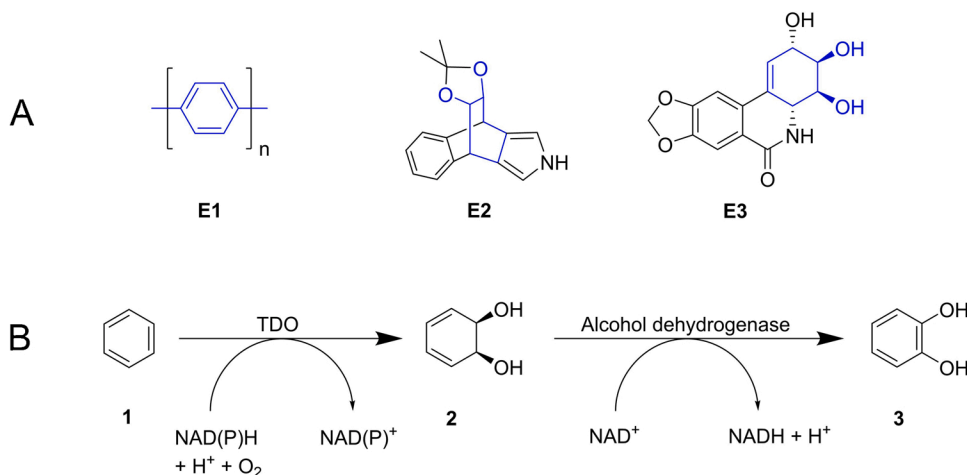
Fig. 1. A: Selected examples of fine chemicals and bioactive compounds of medical relevance harboring the DHC backbone, which is shown depicted in blue; polyphenylene **E1**, precursor of an organic semiconductor **E2**, (+)-lycoricidine **E3**, a drug for cancer treatment, and the anti-diabetic agent (+)-pinitol **E4**. **B:** Observed benzene conversion in *E.coli* BW25113 expressing TDO from *P. putida* F1. Dearomatization-dihydroxylation reaction of benzene **1**, generating DHC **2**, and dehydrogenation reaction, resulting in catechol **3** formation (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

Table 1
Literature compilation reporting the production of DHC by using TDO as biocatalyst. Bacterial strain, working set−up, substrate feeding, and achieved yield of the product are provided. In each case, resting whole cells were employed for benzene biotransformation in shaking flasks, exceptions are marked with superscript letters.

Wild type strain/ recombinant strain (plasmid)	Growth media	Expression media	Reaction volume (mL)	Reaction time (h)	Substrate; benzene (g L^{-1})	Product formation DHC (g L^{-1})	Product isolated yield DHC (g L^{-1})	Reference
P. putida F1	Mineral	Mineral	3	1	0.13	Not reported	Not reported	Gibson et al. (1968)
P. putida 39/D	Mineral	Mineral	100	30	2.63	Not reported	0.23	Gibson et al. (1970)
[a]*P. putida* 11680 mutant strain A	Mineral	Mineral	5000	18	72	13	6.00	Taylor (1995)
[b]*P. putida* UV4	Minimal	Minimal	1500	120	0.58	Not reported	Not reported	Quintana and Dalton (1999)
E. coli JM109 (ISP$_{TOD\alpha}$), hybrid of TDO and BDO	LB	LB	0.5	0.75	0.063	0.09	Not reported	Bagnéris et al. (2005)
E. coli JM101 (pKF6256) hybrid of TDO and BPDO	LB	Minimal	100	48	0.10	Not reported	0.007	Shindo et al. (2005)
[c]*P. putida* KT2442 (pSPM01)	LB	LBG	Not reported	12	4.38	2.68	Not reported	
[a]*P. putida* KT2442 (pSPM01)	LB	LBG	3000	16	140.2	59	Not reported	
P. stutzeri 1317 (pSPM01)	LB	LBG	Not reported	12	4.38	2.13	Not reported	Ouyang et al. (2007)
Aeromonas hydrophilia 4AK4 (pSPM01)	LB	LBG	Not reported	12	4.38	1.17	Not reported	
E. coli JM109 (pSPM01)	LB	LBG	Not reported	12	4.38	0.45	Not reported	
E. coli JM109 (pKST11)	LB	LBG	Not reported	12	4.38	0.53	Not reported	
E. coli BW25113 (pBAD18-TDO)	LB	TB	100	6	3.12	0.85	0.64	This study
E. coli BW25113 Δ*gldA* (pBAD18-TDO)	LB	TB	100	6	3.12	1.76	1.41	

BDO: benzene dioxygenase from *P. putida* ML2.
BPDO: biphenyl dioxygenase from *P. pseudoalcaligenes* KF707.
a; fermenter, **b**; immobilized whole cells in fluidized-bed reactor, **c**; shaking flask, working volume not reported.

Table 2
Bacterial strains employed in this study.

E. coli strain	Genotype	Reference
XL-1 Blue	supE44 hsdR17 recA1 gyrA96 thi relA1 lac F9 [proAB1 laclq lacZDM15 Tn10(Tet^r)]	Stratagene, California, US
BW25113	lacI^q rrnB$_{T14}$ ΔlacZ$_{WJ16}$ hsdR514 ΔaraBAD$_{AH33}$ ΔrhaBAD$_{LD78}$	Datsenko and Wanner (2000)
JM109 (DE3)	endA1 recA1 gyrA96 thi hsdR17 relA1 supE44 λ− Δ(lac-proAB) [F′ traD36 proAB lacI^{q}ZΔM15] lDE3	Studier and Moffatt (1986); Yanisch-Perron et al. (1985)

expression. The cultures were incubated either at 20 or 25 °C at 100 rpm for 20 h. Afterwards, cells were harvested by centrifugation at 6000 × *g* at 4 °C for 20 min and the supernatant was discarded. To prepare the resting cells suspension, the pellet was resuspended to a final concentration of 0.2 g$_{cww}$ mL^{-1} in 0.1 M potassium phosphate buffer (pH 7.4) supplemented with 20 mM glucose monohydrate, for cofactor regeneration.

2.5. Benzene biotransformation using the E. coli BW25113 pBAD18-TDO platform

The resting cells suspension was used to perform benzene

Table 3
Plasmids employed in this study.

Plasmid	Characteristics	Harbored genes	Encoded enzyme	Resistance	Reference
pDTG601	Expression vector	*todC1C2BA*	Toluene dioxygenase (multicomponent system)	Ampicillin	Zylstra and Gibson (1989)
pBAD18	Expression vector	–	–	Ampicillin	Guzman et al. (1995)
pBAD18-TDO	Expression vector	*todC1C2BA*	Toluene dioxygenase (multicomponent system)	Ampicillin	This study

Table 4
Primers employed to generate the pBAD18-TDO expression system.

Primer	Primer name	Amplicon (NCBI Id)	Oligonucleotide primer sequences (5′→ 3′)
1	pBAD18_backbone_Fw	pBAD18	gctagcccaaaaaaacgggtatgg
2	pBAD18_backbone_Rv		ggctgttttggcggatgagagaag
3	todC1_insert_Fw	*todC1*	ccatacccgtttttttgggctagcAGGAGGTAATTCACCatgaatcagaccgacacatcacc
4	todC1_insert_Rv	(5195399)	caatcatGGTGAATTACCTCCTgcttcagcgtgtcgccttcagcgc
5	todC2_insert_Fw	*todC2*	agcAGGAGGTAATTCACCatgattgattcagccaacag
6	todC2_insert_Rv	(5195398)	ccaagtcatGGTGAATTACCTCCTgctctagaagaagaaactgagg
7	todB_insert_Fw	*todB*	ctagagcAGGAGGTAATTCACCatgacttggacatacatattgcg
8	todB_insert_Rv	(5192604)	catGGTGAATTACCTCCTgcttcacttcaactccccg
9	todA_insert_Fw	*todA*	gaagcAGGAGGTAATTCACCatggctacccacgtggcgatc
10	todA_insert_Rv	(5192603)	cttctctcatccgccaaaacagcctcacgttaggtctccttcattcgc

72

biotransformations in 20 mL head space vials with a total reaction volume of 1 mL (liquid to air ratio 1:20). The reaction was started by addition of 20 μL of a 0.5 M benzene solution prepared in ethanol, to a final concentration of 10 mM. Biotransformations were incubated at 30 °C and 180 rpm for 4 h. Reactions were stopped by addition of 0.5 mL of the extraction solvent ethyl acetate. The organic layer was removed and the aqueous phase was extracted a second time with 0.5 mL ethyl acetate. The organic layers were combined and analyzed *via* HPLC-DAD to quantify the achieved product formation of DHC. Additionally, HPLC-ESI-MS analysis was performed to confirm catechol formation.

2.6. Influence of different working parameters on the product formation

In order to ascertain the effect of different working parameters on DHC formation, different conditions to the standard reaction were studied. Such as, 1) reaction time (0, 0.5, 1, 2, 4, 6, and 20 h), 2) substrate cosolvent (paraffin instead of ethanol), 3) pH (6.0, 7.4, and 8.0), 4) benzene concentration (2, 5, 10, and 20 mM), 5) reaction temperature (25, 30, 37, and 45 °C), and 6) benzene feeding fashion to supply 50 mM of the substrate. For each case, the protocol described in section 2.5 above was applied, altering only one of the mentioned conditions at a time.

2.7. Examination of DHC dehydrogenation in E. coli

To confirm that DHC dehydrogenation is due to a native enzymatic reaction, taking place in the host strain, two different approaches were tested; 1) In addition to the *E. coli* BW25113 strain, we also used the JM109(DE3) strain. In both cases, strains were not harboring the pBAD18-TDO system. Cell cultures and resting cells were prepared as described in section 2.4, omitting the induction step. 2) *E. coli* BW25113 strain resting cells were disrupted in a high pressure homogenizer (Emulsiflex C5, Avestin, Ottawa, CA), and used to obtain cell free lysates. To denaturize the protein pool, the lysate was heat-shocked at 50 and 95 °C and used for biotransformations. To examine DHC dehydrogenation reaction, DHC instead of benzene was applied as substrate for both approaches (1 and 2), according to section 2.5, with subsequent analysis *via* HPLC-DAD.

2.8. KEIO-collection-based screening approach to detect DHC consumption

For identification of the native dehydrogenase(s) from *E. coli* BW25113 responsible for the DHC conversion, 88 different *E. coli* BW25113 knock-out strains from the KEIO collection were examined. The strains were precultured in 96-deep well plates, containing 0.5 mL of LB media supplemented with 50 μg mL^{-1} of kanamycin, using glycerol stocks for inoculation. The plates were sealed with Breathe-EASIER sealing membranes for multiwell plates (Sigma, St. Louis, US). Incubation was performed for 20 h at 37 °C and 800 rpm in a microplate shaker (VWR International, Radnor, US). To assess DHC consumption, cell cultures were grown in 96-deep well plates, containing 1.0 mL TB medium, 2.5 % (v/v) of preculture, 50 μg mL^{-1} kanamycin, and incubated at 37 °C and 800 rpm for 20 h. Cells were harvested by centrifugation at 4000 x *g* and 4 °C for 20 min, and the supernatant was discarded. Biotransformations were performed in 96-deep well plates by resuspending the cell pellets in 490 μL of 0.1 M potassium phosphate buffer (pH 7.4) supplemented with 20 mM glucose. A volume of 10 μL of the 0.1 M DHC solution (in ethanol, final concentration 2 mM), was added to each different strain. Biotransformation plates were incubated at 30 °C and 800 rpm for 20 h. Biotransformations were stopped by addition of 1:1 (v/v) ethyl acetate and centrifuged. Afterwards, the ethyl acetate layer was separated and analyzed by HPLC-DAD to quantify DHC consumption.

2.9. Biosynthesis of DHC using the enhanced pBAD18-TDO E. coli BW25113 platform

In order to obtain sufficient amounts of DHC, a semi-preparative scale reaction was performed. The *E. coli* strains BW25113 and BW25113 Δ*gldA* were employed, as expression hosts, to produce the resting cells harboring the pBAD18-TDO system. Protein expression and resting cells suspension preparation were performed according to the conditions described in section 2.4 above, but with an expression temperature of 20 °C. Regarding the set-up for benzene biotransformation, the protocol described in section 2.5 above was applied, implementing some modifications. The reaction temperature was decreased to 25 °C, the total reaction volume was up-scaled to 100 mL in 2 L Schott flasks as reaction vessel. The total substrate addition was 40 mM, using a feeding approach. Substrate feeding was performed in three steps; first addition was 20 mM to start the reaction and additional 10 mM were fed at 2 and 4 h after starting the reaction. The total reaction time was 6 h. The whole reaction volume was extracted five times with ethyl acetate (1:1, v/v), and the organic layers were combined. After *in vacuo* evaporation of the organic solvent, the resulting product mixture was purified by silica gel column chromatography. To prevent DHC dehydrogenation, silica gel (60 M, 0.04−0.63 mm) was prepared with sodium hydrogen carbonate solution according to Nagy and colleagues (Nagy et al., 2009). An ethyl acetate:cyclohexane elution gradient (5:1 to 1:0) was employed for fractionation. The fractions were analyzed by thin layer chromatography. Key fractions, containing DHC were additionally analyzed by HPLC-DAD and HPLC-ESI-MS. The purified product was submitted to ^{1}H- and ^{13}C-NMR for structural elucidation. Purity assessment was performed by HPLC-DAD.

2.10. Analytical methods

2.10.1. HPLC-DAD set-up for DHC and catechol quantification

For benzene biotransformations, extracted samples were analyzed by high-pressure liquid chromatography (HPLC). An Agilent 1260 Infinity II system (Santa Clara, US), equipped with a C18-column (Agilent Eclipse XDB-C18, 5 μm, 4.6 × 150 mm, Santa Clara, US) and a diode array detector (Agilent 1260 Infinity II DAD HS, Santa Clara, US) was operated isothermally at 30 °C. Measurements were run at a flow rate of 0.6 mL min^{-1}. The mobile phase was water/acetonitrile with a linear gradient of; t = 0 min, 35/65 (v/v); t = 5.5 min, 10/90 (v/v); t = 6.0 min, 35/65 (v/v); t = 10.0 min, 35/65 (v/v). Peak areas were measured by the integrator and transformed into concentration using the correspondent standard curves of DHC and catechol (Fig. S11). The employed wavelength to detect DHC and catechol were 262 and 210 nm, respectively.

2.10.2. HPLC-ESI-MS set-up for catechol identification

To fully confirm the catechol generation during benzene transformations, a HPLC-ESI-MS was established for catechol ion detection. An Agilent 6130 Quadrupole LC System (Santa Clara, US), was operated using identical HPLC method and column described in section 2.10.1, above. The ESI-MS was run using the following parameters: API-ES spray chamber with a drying gas flow of 12 L min^{-1}, nebulizer pressure of 50 psi, drying gas temperature 350 °C and a capillary voltage of 3500 V positive and negative. Mass detection was performed in negative SCAN mode from 50 to 350 *m/z*, the obtained MS spectra for the catechol standard and for the catechol generated during biotransformations are shown in figures S13 and S14, respectively.

2.11. Nuclear magnetic resonance for structural characterization of DHC

The purity of the biosynthesized DHC was assessed by ^{1}H- and ^{13}C-NMR. Spectra were recorded on a Bruker Avance 500 spectrometer operating at 500.15 and 125.76 MHz, respectively. Spectra were recorded using D$_2$O as solvent at room temperature. The chemical shifts (δ) are reported in parts per million (ppm) relative to the standard

tetramethylsilane (TMS, $\delta = 0$).

3. Results and discussion

3.1. Establishment of the pBAD18-TDO system for DHC biosynthesis

The initial aim of this work was to achieve a gram per liter synthesis and isolation of DHC. We pursued this task by constructing a new expression system for the TDO platform. We chose the arabinose inducible pBAD18 plasmid to harbor the *tod* genes encoding the TDO system and to drive protein expression. The benefits of the pBAD18-TDO system are; low cost of the inductor (L-arabinose), tight control of the induction, and enhancement of the protein solubility in *E. coli* (Guzman et al., 1995). We selected *E. coli* BW25113 as host strain, since it is lacking the *araBAD* operon involved in arabinose catabolism (Datsenko and Wanner, 2000). In order to test and evaluate the efficiency of the generated TDO platform, we established an initial protocol to generate the resting cells expressing the TDO system and a standard reaction working setup using naphthalene as model substrate to assess activity (Wissner et al., 2020). As a result, the pBAD18-TDO system successfully transformed the total concentration (10 mM) of the supplied naphthalene into (1*R*,2*S*)-1,2-dihydronaphtalene-1,2-diol in 1 h (data not shown). Nevertheless, when using benzene (10 mM) as substrate, a low DHC formation (0.5 mM) was detected after a biotransformation time of 20 h. In addition, we observed the generation of large quantities (6.2 mM) of an unexpected byproduct that was identified as catechol by HPLC-ESI-MS (Fig. S14).

3.2. Influence of different working parameters on DHC formation

To investigate the reason behind such low DHC formation and the generation of the unexpected byproduct (catechol), we performed a time-lapse experiment. Biotransformation reactions were stopped at different reaction times and analyzed regarding DHC and catechol concentrations (Fig. 2A). Results showed that the highest DHC formation of 2.1 mM is achieved after 4 h, but this concentration substantially drops in the following 16 h, given that only 0.6 mM can be detected after 20 h. Along with the substantial reduction of DHC a catechol buildup was detected, displaying a gradual inverse correlation, in concentration, through the time, resulting in 7.1 mM catechol after 20 h. This suggests that DHC might be converted to catechol. Subsequently, we wanted to assess the influence of the employed temperature for protein expression, thus we decreased the expression temperature from 25 to 20 °C. We performed a second time-lapse experiment, but this time we employed resting cells harboring the pBAD18-TDO system that were expressed at 20 °C. Once again, biotransformation reactions were analyzed at different time points regarding product and byproduct formation (Fig. 2B). Results showed a significant improvement in terms of DHC formation, up to 4.6 mM after 4 h, and additionally a favorable

reduction of the undesired byproduct. Nevertheless, after a reaction time of 20 h, the concentration of the generated DHC gradually decreases to 2.8 mM, while catechol increases in concentration to 5.7 mM. This finding also supports our assumption that catechol is generated from DHC. The data of the time-lapse experiments suggest that a reaction time of 4 h, in addition to an expression temperature of 20 °C, favors DHC formation, while abating considerably byproduct generation. This might be explained by the fact that the TDO was better expressed and that the reaction responsible for the catechol formation was slowed down. The striking formation of catechol from benzene in *E. coli* expressing the TDO is so far not described in literature.

Since protein expression temperature showed to have a noticeable positive influence on DHC generation, we wondered whether varying working parameters, as described in section 2.6, also might substantially favor the final product formation (Fig. 3). We started by examining the influence of liquid paraffin as cosolvent, since it can reduce cell toxicity and volatility of benzene (Ouyang et al., 2007). Results showed that using paraffin instead of the regular cosolvent ethanol (Halder et al., 2018), had no significant influence on the DHC and catechol formation. Considering that paraffin showed not a beneficial influence on product formation and might complicate the downstream extraction and purification procedure (Ouyang et al., 2007), ethanol was continued to be used as cosolvent.

We also evaluated the effect of lower and higher pH values with pH 6.0 and 8.0. At pH 6.0 a decrease of catechol formation was observed, while DHC concentration remained at the same amount as in the standard condition. However, formation of phenol was detected due to DHC dehydration, which was neither observed at pH 7.4, nor at 8.0. In contrast, at pH 8.0 catechol generation was more favored. According to our findings and the literature (Ouyang et al., 2007; Quintana and Dalton, 1999), we continued with pH 7.4 for further experiments.

We observed that the higher the benzene concentration, the higher the overall product formation and the higher the DHC/catechol ratio, which means 8.9 mM of product when 40 mM of benzene is supplied. This suggest that benzene concentration have a positive influence on DHC formation, but only a minor one, on catechol generation.

By altering the reaction temperature we could assess that with a reaction temperature of 37 °C, the total conversion is fostered up to 6.5 mM, nevertheless the product distribution goes more in favor to the undesired byproduct catechol. At 45 °C the total conversion is reduced to half compared to standard conditions, most likely due to protein denaturation. At the lowest reaction temperature 25 °C, a higher DHC concentration was achieved (2.5 mM). Thus, for the further reactions, including for the preparative scale, 25 °C were used as reaction temperature.

Regarding the influence of different working parameters on product formation, results showed that in addition to the most suited reaction time (4 h), protein expression temperature, reaction temperature, and the supplied substrate concentration also positively influence DHC
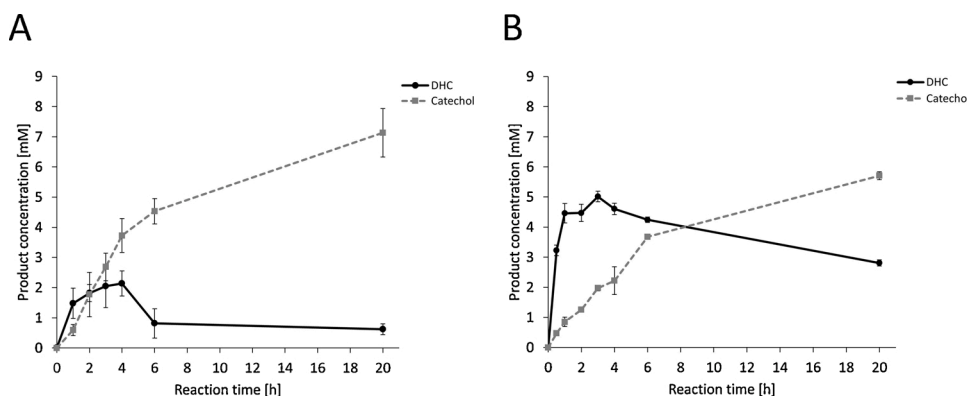

Fig. 2. Influence of protein expression temperature on the conversion of benzene to DHC (black) and secondary reaction generating catechol (grey dotted). **A**: TDO expressed at 25 °C for 20 h. **B**: TDO expressed at 20 °C for 20 h. For both cases, biotransformations were performed with *E. coli* BW25113 resting cells harboring the pBAD18-TDO system (0.2 g_{cww} mL^{-1}), 0.1 M potassium phosphate buffer (pH 7.4), 20 mM glucose, and 10 mM benzene, at 30 °C and 180 rpm for 20 h. Each point shows the mean value of three biological triplicates and their corresponding standard deviations. Mass balance at any given time point might not be fully closed due to benzene volatility.
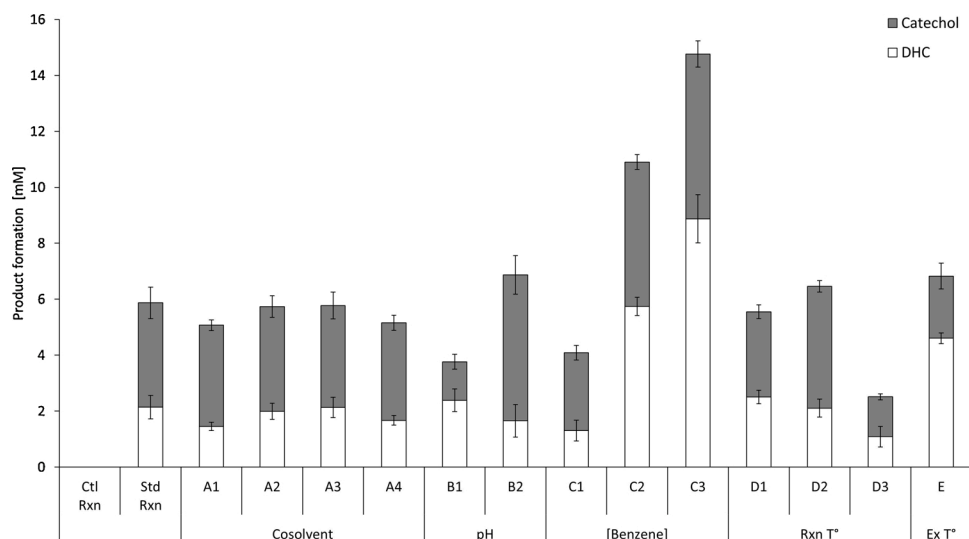
Fig. 3. Influence of different working parameters on DHC and catechol formation and distribution after 4 h reaction time. The standard reaction condition was implemented as follows; *E. coli* BW25113 whole cells harboring pBAD18-TDO system, protein expressed at 25 °C for 20 h, 0.2 g_{cww} mL^{-1} cells suspended in 0.1 M potassium phosphate buffer (pH 7.4) supplemented with 20 mM glucose and 10 mM benzene, were used for biotransformations performed at 30 °C. Each condition followed the same setup as in the standard reaction by modifying, one of the following parameters separately; Ctl Rxn = control reaction, buffer system with substrate; Std Rxn = standard reaction (as stated above); A1-A4 = addition of cosolvent (A1 = direct addition of benzene; A2 = benzene:paraffin 1:3; A3 = benzene:paraffin 1:22; A2 = benzene:paraffin 1:222); B1-B2 = reaction buffer pH (B1 = pH 6.0; B2 = pH 8.0); C1-C3 = benzene concentration (C1 = 5 mM; C2 = 20 mM; C3 = 40 mM); D1-D3 = reaction temperature (D1 = 25 °C; D2 = 37 °C; D3 = 45 °C); E = expression temperature (20 °C). Each bar is showing the mean value of three biological triplicates and the corresponding standard deviations.

formation.

By combining a protein expression temperature of 20 °C, a reaction time of 4 h, a reaction temperature of 25 °C, and a benzene concentration of 10 mM, we accomplished to foster DHC formation up to 4.9 mM and catechol formation was reduced to 0.9 mM (Fig. S12). This, in comparison to the standard working set up, represents a 2.3-fold improvement of DHC formation and a 4.1-fold decrease in catechol formation.

Results showed that adding higher benzene concentrations than in the standard reaction actually favored DHC formation, nonetheless, a high concentration of benzene has been shown to be detrimental for the resting cells due to its toxicity (Ouyang et al., 2007). Considering our results, we assumed that a substrate feeding strategy might foster product formation, avoid toxicity issues, and hamper at some degree the secondary reaction to catechol. We supplied different amounts of benzene at different time points under the enhanced working conditions set up. The most favorable substrate concentration to generate DHC was 40 mM in the feeding strategy fashion. Whereby, the substrate addition was split in three steps; 20 mM when the reaction was started, followed by 10 mM after 2 h, and 10 mM after 4 h, with a total reaction time of 6 h (Fig. S1). Therefore, we employed this set up for semi-preparative application. Contrary to our expectations, biotransformation reactions fed with benzene concentrations higher than 40 mM did not showed an increase in DHC formation. The reason behind can be conferred to the high benzene toxicity to bacterial cells (Ouyang et al., 2007).

3.3. Semi-preparative biotransformation with BW25113 pBAD18-TDO

Semi-preparative biotransformation at 100 mL scale resulted in 7.6 mM DHC and 2.2 mM catechol, 19 and 6% product formation, respectively. This DHC concentration represents 0.85 g L^{-1}, which is comparable to shaking flask experiments reported by Ouyang and colleagues (Ouyang et al., 2007). After chromatographic column purification, an isolated yield of 0.64 g L^{-1} was achieved. The structural identity of the purified product was confirmed *via* ^{1}H- and ^{13}C-NMR (Figs. S15 and S16, respectively). Assessing product purity was important since it would be used as substrate to perform the experiments regarding DHC dehydrogenation. Analysis *via* HPLC-DAD confirmed high DHC purity and the absence of catechol (Fig. S10). Thus, the produced molecule appeared to be suitable as standard and as substrate for the

characterization of the downstream reaction from DHC to catechol.

Description of the biosynthesized *cis*-1,2-dihydrocatechol: Colorless crystals. ^{1}H-NMR (D$_2$O, 500.15 MHz) δ = 4.28 (s, 2 H), 5.95 (m, 2 H), 6.06 (m, 2 H) ppm. ^{13}C-NMR (D$_2$O, 125.76 MHz) δ = 70.10, 128.02, 131.24 ppm.

3.4. Investigations on DHC dehydrogenation in E. coli

As we mentioned earlier, obtaining catechol in our biotransformations was completely unanticipated. The results described in Sections 3.1 and 3.2 suggest that catechol is a dehydration product generated from DHC. Catechol formation increases over time reaching its maximum between six and 20 h; nonetheless, its formation can be reduced when lower expression temperatures are applied during protein expression. Besides, catechol formation was slightly decreased by altering working parameters such as reaction temperature, reaction time and substrate concentration. However, despite implementing the enhanced setup, catechol was still formed.

3.4.1. Examinations on DHC dehydrogenation in E. coli

In order to enlighten the source of DHC dehydrogenation we started by ruling out whether the expression host itself was responsible such reaction. We prepared wild type *E. coli* BW25113 resting cells, without the pBAD18-TDO system, for the biotransformation of DHC as substrate. After 1 h reaction, *E. coli* BW25113 resting cells converted 70 % of the DHC into catechol, while in the control reaction lacking resting cells no conversion was detected (Fig. S2). Therefore, the dehydration reaction seems to be performed by one or more enzymes from the enzymatic pool of *E. coli* BW25113. We wanted to assess whether this reaction was taking place merely in the BW25113 genetic background. We repeated the same experiment with the *E. coli* JM109 (DE3) strain. Indeed, the findings were the same as with BW25113, catechol was formed in the same amount and in the same reaction time (Fig. S2). Since we used resting cells lacking the pBAD18-TDO system and DHC as substrate, we could rule out that the dehydrogenation reaction was either due to the expressed protein complex, or to the substrate benzene. Considering that *E. coli* BW25113 resting cells are directly driving DHC dehydrogenation, we wanted to determine whether the reaction was coming from the cellular cytosolic fraction. We prepared heat-shocked cell free extracts, as described in section 2.7, and we followed the same protocol as for

whole cells but, this time the reactions were performed for 20 h to resemble the examined time window in the time-lapse experiments. According to figure S3, the control reaction showed no DHC conversion, while the non-denatured cell free extract produced catechol, as observed in whole cells and in the enhanced pBAD18-TDO platform. The catechol formation decreased by 15 % with cell free extracts that were heat-shocked at 50 °C in comparison with the non-denatured one, whereas it was completely prevented when cell extracts were denatured at 95 °C (Fig. S3). These findings, all together, confirm that cytosolic pool from *E. coli* BW25113 was harboring the responsible enzyme(s) for DHC dehydrogenation reaction. In an attempt to identify such enzyme(s), we search in the literature for the reported enzymes capable to dehydrogenate DHC (Table S2). In *P. putida* F1, which is the natural host of the TDO system, the enzyme *cis*-toluene dihydrodiol dehydrogenase TodD is the responsible for converting DHC into catechol (Table S4) (Zylstra and Gibson, 1989). Regarding known enzymes from *E. coli* displaying such activity, the enzyme 2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase EntA, is the only one so far reported (Table S4) (Sakaitani et al., 1990). We use TodD sequence as query for searching homologs in *E. coli*. The analysis showed a 3-phenylpropionate-dihydrodiol dehydrogenase HcaB as the closest hit, in terms of sequence identity (43 %), with respect with TodD. While, an EntA alignment analysis, using TodD sequence as query, showed only a 23 % identity. To corroborate whether HcaB or EntA were involved in DHC degradation in our established platform we used *E. coli* BW25113 Δ*hcaB* or Δ*entA* cells from the KEIO collection.

3.4.2. KEIO-collection-based screening approach to detect DHC consumption

The KEIO collection (Baba et al., 2006), was generated by using *E. coli* BW25113 as genetic background and includes every single gene deletion of all non-essential genes. Thus, this mutant collection provides a resource for the analysis of gene-protein associated functions. We tested *E. coli* BW25113 Δ*hcaB* or Δ*entA* resting cells activity towards DHC, as described in section 2.8, under the assumption that no DHC conversion into catechol would indicate that the deleted gene was encoding the protein responsible for the undesired reaction. Surprisingly, both of the tested mutant strains were capable to generate catechol, showing that none of them had a noticeable influence in the reaction (Fig. S4). We hypothesized that in *E. coli* BW25113, an alcohol dehydrogenase type enzyme(s) should be involved in catechol formation. Once again, we used the KEIO collection as a resource to investigate the connection between the *E. coli* dehydrogenases pool and catechol generation. We analyzed 88 different dehydrogenase knock-out strains (Table S1), (protocol described in section 2.8), regarding their capability to degrade DHC. This resulted in the identification of the *E. coli* BW25113 Δ*gldA* strain as incapable to fully degrade our valuable molecule, since only a small amount of the substrate (5%), was converted. Such findings revealed that the protein glycerol dehydrogenase GldA was the main responsible enzyme for DHC degradation into catechol (Table S2). Since the result was quite striking, we searched for protein homologs using GldA sequence as query, resulting in the identification of *cis*-benzene dihydrodiol dehydrogenase BedD with an identity of 73 %. BedD naturally catalyzed the dehydrogenation of DHC into catechol in the benzene degradation pathway from *P. putida* ML2 (Table S4). In addition, we also performed a protein sequence alignment between BedD and TodD from *P. putida* F1, resulting in a surprising very low identity of only 17 %, despite both enzymes are capable to perform the same reaction. Fong and colleagues, previously reported the homology between BedD and GldA from *E. coli* (Fong et al., 1996). It is also known, that GldA displays a strong thermostability at 60 °C (Piattoni et al., 2013), which explains why a 50 °C heat-shock produced only a minor impact on DHC conversion. Since GldA is an alcohol dehydrogenase, and we employed ethanol as cosolvent to supply either benzene or DHC to the reactions, we wondered whether the GldA activity was induced by ethanol. We performed reactions supplying only benzene alone to *E. coli* BW25113 resting cells. Despite the cosolvent ethanol was

absent, catechol was formed in the same amount, demonstrating that ethanol is not inducing the dehydrogenation activity (Fig. 3, A1). To fully corroborate GldA role in DHC dehydrogenation, we assessed the presence and the lack of *gldA* gene in *E. coli* BW25113 and in BW25113 Δ*gldA* strains, respectively. We obtained the genomic DNA of both strains and performed PCR reactions, employing primers capable to amplify either the *gldA* gene or the inserted kanamycin cassette replacing the gldA gene (Supplementary section 3, Tables S5, S6 and S7). The electrophoretic analysis of the amplicons (Fig. S5), confirmed that BW25113 strain was harboring the *gldA* gene (also verified by Sanger sequencing). In the case of the BW25113 *gldA* mutant strain, the gene was effectively replaced by the kanamycin cassette as expected. Therefore, there is no doubt about our initial inferences regarding the role of GldA as responsible for catechol generation based on the obtained results described in figure S4. Once we identified the main enzyme involved in DHC degradation, the next step was transferring the pBAD18-TDO system into the *E. coli* BW25113 Δ*gldA* strain. We followed the product distribution over time, resembling the time lapse experiment described in section 3.1, including the comparison between two different expression temperatures. The transfer of the pBAD18-TDO system, into the *E. coli* BW25113 Δ*gldA* strain, resulted in a pronounced enhancement in terms of DHC production, since the dehydration reaction of the product was nearly entirely prevented (Fig. 4). The inferred total conversion of benzene was in the same range (75–88 %) after a reaction time of 20 h, independently of the host strain or the protein expression temperature. However, the product distribution favoring the ratio to the desired product, DHC, was more prominent when the protein was expressed at 20 °C and with *E. coli* BW25113 Δ*gldA* as host strain (Fig. 4B). Under these conditions, the maximum DHC formation (9.0 mM) was reached after 4 h reaction and stayed practically constant over time, up to the last measurement performed at 20 h. Regarding catechol build up, only minor traces were detected (0.1 mM), the explanation for this negligible DHC dehydrogenation might be because the presence of other alcohol dehydrogenases showing a very slender affinity for this substrate. This is in agreement with our findings showed in figure S4, indicating that GldA was the main responsible for the dehydrogenation reaction but, not the only dehydrogenase capable to use DHC as substrate to generate catechol. In view of the cooperative benefits on moving the pBAD18-TDO platform into the *E. coli* BW25113 Δ*gldA* chassis, in combination with a protein expression performed at 20 °C, the next obvious step was applying, in addition, the identified working parameters positively influencing DHC formation.

3.4.3. Biosynthesis of DHC using the enhanced pBAD18-TDO E. coli BW25113 platform

As it was mentioned in point 3.3, we accomplished the production of 0.85 g L^{-1} of isolated DHC. Nonetheless, the genuine challenge was still to achieve the gram per liter synthesis. Therefore, we performed a new semi-preparative 100 mL scale biotransformation of benzene (Section 2.9). The aim was to produce more product, but also to assess the additive or even synergistic effects of integrating the most prominent identified factors positively influencing DHC formation. We combined the pBAD18-TDO platform along with the *E. coli* BW25113 Δ*gldA* chassis, protein expression performed at 20 °C, and the identified working parameters favoring DHC formation. This time DHC formation was fostered up to 15.7 mM (39 % product formation), which represents 1.76 g L^{-1}. This denotes a more than twofold higher product concentration in comparison to the previous biotransformation (Section 3.3), with wild type *E. coli* BW25113 cells. It is noteworthy to mention, that not even traces of catechol formation were detected. After chromatographic column separation, 1.41 g L^{-1} yield of pure product was isolated, fulfilling our goal of producing DHC in gram per liter scale. The structural identity of the molecule was confirmed by ^{1}H- and ^{13}C-NMR. To our knowledge this is the highest DHC yield, obtained by a recombinant chassis at semi-preparative scale (Table 1).
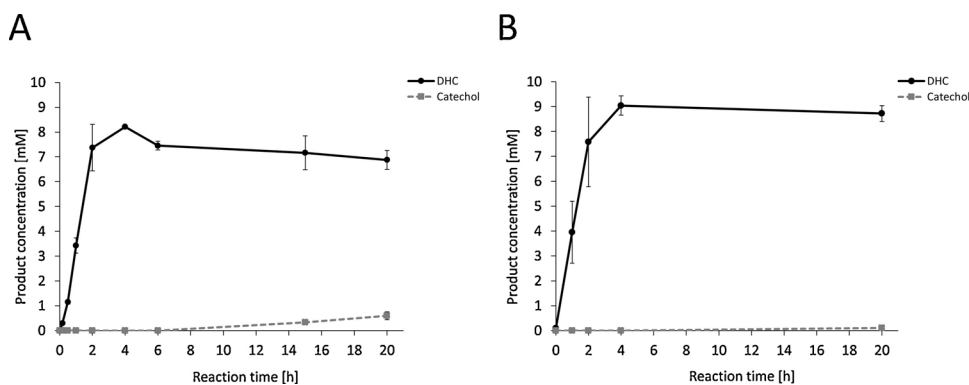
A

B

3.4.4. Comparison of BW25113 wild type and ΔgldA for toluene

After identifying the glycerol dehydrogenase GldA as the main responsible for DHC dehydrogenation, we found a 73 % sequence identity with the benzene dehydrogenase BedD from *P. putida* ML2. According to Fong and colleagues (Fong et al., 1996; Fong and Tan, 1999), the substrate scope of BedD includes, in addition to DHC; the triol glycerol, the diol 1,2-propanediol, and the *cis*-diols; *cis*-1,2-cyclohexanediol, *cis*-toluene dihydrodiol, and *cis*-naphthalene dihydrodiol (Table S3). So far, the substrate scope of GldA towards *cis*-diols has never been explored. In our journey directed to enhance the TDO platform for the production for DHC, we faced the problem of an unforeseen reaction dehydrating our product to generate catechol. Our investigations on DHC dehydrogenation in *E.coli*, led to the discovery of GldA as an alcohol dehydrogenase displaying an impressive affinity towards the substrate DHC. Thus, to explore further the substrate scope of GldA, we evaluated the conversion of toluene (10 mM), a larger substrate that benzene, in the enhanced TDO platform installed either in wild type BW25113 or in the *E. coli* BW25113 Δ*gldA* chassis (Supplementary section 4). Since toluene is the natural substrate of TDO, we observed a full conversion of the substrate in 1 h, generating *cis*-1,2-dihydro-3-methylcatechol. We assumed that in the case this molecule was included in the substrate scope of GldA, a noticeable formation of 3-methylcatechol might be observed (Figs. S6 and S8). Indeed, results showed that along with *cis*-1,2-dihydro-3-methylcatechol generation, also 3-methylcatechol was build up. To investigate the product distribution over time, we perform a time lapse experiment with both strains as described in section 2.5, following 3-methylcatechol generation (Fig S6). In both cases, full conversion of toluene was reached after 1 h, but only minor amounts of 3-methylcatechol were formed over time (Figs. S7 and S9). After a reaction time of 20 h, results showed only a partial reduction in 3-methylcatechol formation in the *E. coli* BW25113 Δ*gldA* strain in comparison with the wild type (0.28 mM and 0.55 mM, respectively). This differs with catechol formation, which was nearly totally prevented by using the *E. coli* BW25113 Δ*gldA* genetic background. Since, the dehydrogenation reaction was only partially reduced in the strain lacking the *gldA* gene, it seems that other alcohol dehydrogenases besides GldA are capable to generate 3-methylcatechol in *E. coli* BW25113 strain. In addition, due to the low 3-methylcatechol formation, it can be assumed, that GldA displays a lower affinity towards *cis*-1,2-dihydro-3-methylcatechol than for DHC. These results are in agreement with the findings of Bagnéris and colleagues. They observed catechol formation as a degradation product of DHC, but they did not detected signals of degradation of *cis*-1,2-dihydro-3-methylcatechol during the conversion of toluene by TDO in *E. coli* (Bagnéris et al., 2005). However, they assumed that the degradation of DHC to catechol was an artefact from thermal decomposition during the GC–MS analysis. This supposition is questionable since the dehydrogenation of DHC to catechol seems quite unlikely, due to the favoured dehydration to phenol. During GC–MS analysis of DHC we could only detect phenol as degradation product (data not shown).

4. Conclusion

In this work we exploited TDO as biocatalyst to perform the direct dearomatizing dihydroxylation of benzene to produce DHC in a one-step reaction. The following specific points were achieved; establishment of a new TDO expression system designed for a recombinant *E. coli* chassis, identification of key working parameters for enhanced product formation, discovery and elucidation of an unforeseen reaction degrading DHC into catechol, and identification of the appropriate genetic background to prevent DHC dehydrogenation, *E. coli* BW25113 Δ*gldA*. We combined the most prominent factors such as protein expression temperature, substrate concentration, substrate feeding fashion, and appropriate host cell, to foster DHC formation. As a result, the enhanced platform made feasible our goal to achieve a gram per liter production of DHC. Moreover, further studies regarding the substrate spectrum of GldA are currently under preparation. In summary, our results indicate that *E. coli* BW25113 Δ*gldA* strain harboring pBAD18-TDO system is a suitable platform for the production of the valuable compound DHC at gram scale. This work represents a starting setup for the further establishment of the TDO platform at preparative scale for actual industrial applications.

CRediT authorship contribution statement

Julian L. Wissner: Conceptualization, Investigation, Methodology, Visualization, Writing - original draft. **Julian Ludwig:** Investigation, Methodology, Visualization. **Wendy Escobedo-Hinojosa:** Methodology, Validation, Writing - review & editing. **Bernhard Hauer:** Funding acquisition, Supervision, Writing - review & editing.

Declaration of Competing Interest

The authors report no declarations of interest.

Acknowledgements

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.jbiotec.2020.09.012.

References

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., Mori, H., 2006. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol. Syst. Biol. 2 https://doi.org/10.1038/msb4100050.

Bagnéris, C., Cammack, R., Mason, J.R., Bagne, C., 2005. Subtle difference between benzene and toluene dioxygenases of Pseudomonas putida. Appl. Environ. Microbiol. 71, 1570–1580. https://doi.org/10.1128/AEM.71.3.1570.

Ballard, D.G.H., Courtis, A., Shirley, I.M., Taylor, S.C., 1983. A biotech route to polyphenylene. J. Chem. Soc. Chem. Commun. 954–955. https://doi.org/10.1039/c39830000954.

Datsenko, K.A., Wanner, B.L., 2000. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc. Natl. Acad. Sci. U. S. A. 97, 6640–6645. https://doi.org/10.1073/pnas.120163297.

Fong, K.P.Y., Tan, H.M., 1999. Characterization of a novel cis-benzene dihydrodiol dehydrogenase from Pseudomonas putida ML2. FEBS Lett. 451, 5–9. https://doi.org/10.1016/S0014-5793(99)00520-7.

Fong, K.P.Y., Goh, C.B.H., Tan, H.M., 1996. Characterization and expression of the plasmid-borne bedD gene from Pseudomonas putida ML2, which codes for a NAD+-dependent cis-benzene dihydrodiol dehydrogenase. J. Bacteriol. 178, 5592–5601. https://doi.org/10.1128/jb.178.19.5592-5601.1996.

Gibson, D.G., 2011. Enzymatic assembly of overlapping DNA fragments. Methods Enzymol. 498, 349–361. https://doi.org/10.1016/B978-0-12-385120-8.00015-2.

Gibson, D.T., Koch, J.R., Kallio, R.E., 1968. Oxidative degradation of aromatic hydrocarbons by microorganisms. I. enzymatic formation of catechol from benzene. Biochemistry 7, 2653–2662. https://doi.org/10.1021/bi00847a031.

Gibson, D.T., Cardini, G.E., Maseles, F.C., Kallio, R.E., 1970. Incorporation of Oxygen-18 into benzene by Pseudomonas putida. Biochemistry 9, 1631–1635. https://doi.org/10.1021/bi00809a024.

Guzman, L.-M., Belin, D., Carson, M., Beckwith, J., 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose P(BAD) promoter. J. Bacteriol. 177, 4121–4130. https://doi.org/10.1128/jb.177.14.4121-4130.1995.

Halder, J.M., Nestl, B.M., Hauer, B., 2018. Semirational Engineering of the Naphthalene Dioxygenase from Pseudomonas sp. NCIB 9816-4 towards Selective Asymmetric Dihydroxylation. ChemCatChem 10, 178–182. https://doi.org/10.1002/cctc.201701262.

Hanahan, D., 1983. Studies on transformation of Escherichia coli with plasmids. J. Mol. Biol. 166, 557–580. https://doi.org/10.1016/S0022-2836(83)80284-8.

Hudlicky, T., Olivo, H.F., 1992. A short synthesis of (+)-lycoricidine. J. Am. Chem. Soc. 114, 9694–9696. https://doi.org/10.1021/ja00050a079.

Ley, S.V., Sternfeld, F., Taylor, S., 1987. Microbial oxidation in synthesis: a six step perparation of (+)-pinitol from benzene. Tetrahedron Lett. 28, 225–226. https://doi.org/10.1016/S0040-4039(00)95692-2.

Masumoto, A., Kikuchi, T., Ono, N., Uno, H., Nakashima, H., 2011. COMPOUND AND METHOD OF PRODUCING ORGANIC SEMICONDUCTOR DEVICE. US007928221B2.

Nagy, V., Agócs, A., Turcsi, E., Deli, J., 2009. Isolation and purification of acid-labile carotenoid 5,6-epoxides on modified silica gels. Phytochem. Anal. 20, 143–148. https://doi.org/10.1002/pca.1109.

Nakajima, M., Tomida, I., Takei, S., 1959. Zur Chemie Des Benzolglykols, IV. Darstellung voncis-Benzolglykol, stereoisomeren mucondialdehyden und Kondurit-F. Chem. Ber. 92, 163–172. https://doi.org/10.1002/cber.19590920119.

Ouyang, S.-P., Sun, S.-Y., Liu, Q., Chen, J., Chen, G.-Q., 2007. Microbial transformation of benzene to cis-3,5-cyclohexadien-1,2-diols by recombinant bacteria harboring toluene dioxygenase gene tod. Appl. Microbiol. Biotechnol. 74, 43–49. https://doi.org/10.1007/s00253-006-0637-6.

Paddock, V.L., Phipps, R.J., Conde-Angulo, A., Blanco-Martin, A., Giró-Mañas, C., Martin, L.J., White, A.J.P., Spivey, A.C., 2011. (±)- Trans, cis -4-Hydroxy-5,6-di- O -isopropylidenecyclohex-2-ene-1-one: synthesis and Facile dimerization to decahydrodibenzofurans. J. Org. Chem. 76, 1483–1486. https://doi.org/10.1021/jo102314w.

Parales, R.E., Resnick, S.M., 2006. Biocatalysis in the pharmaceutical and biotechnology industries. Biocatalysis in the Pharmaceutical and Biotechnology Industries. CRC Press. https://doi.org/10.1201/9781420019377.

Piattoni, C.V., Figueroa, C.M., Asención Diez, M.D., Parcerisa, I.L., Antuña, S., Comelli, R. A., Guerrero, S.A., Beccaria, A.J., Iglesias, A.Á., 2013. Production and characterization of Escherichia coli glycerol dehydrogenase as a tool for glycerol recycling. Process Biochem. 48, 406–412. https://doi.org/10.1016/j.procbio.2013.01.011.

Quintana, M.G., Dalton, H., 1999. Biotransformation of aromatic compounds by immobilized bacterial strains in barium alginate beads. Enzyme Microb. Technol. 24, 232–236. https://doi.org/10.1016/S0141-0229(98)00116-1.

Sakaitani, M., Rusnak, F., Quinn, N.R., Tu, C., Frigo, T.B., Berchtold, G.A., Walsh, C.T., 1990. Mechanistic studies on trans-2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase (Ent A) in the biosynthesis of the iron chelator enterobactin. Biochemistry 29, 6789–6798. https://doi.org/10.1021/bi00481a006.

Shindo, K., Nakamura, R., Osawa, A., Kagami, O., Kanoh, K., Furukawa, K., Misawa, N., 2005. Biocatalytic synthesis of monocyclic arene-dihydrodiols and -diols by Escherichia coli cells expressing hybrid toluene/biphenyl dioxygenase and dihydrodiol dehydrogenase genes. J. Mol. Catal. B Enzym. 35, 134–141. https://doi.org/10.1016/j.molcatb.2005.06.010.

Studier, F.W., Moffatt, B.A., 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J. Mol. Biol. 189, 113–130. https://doi.org/10.1016/0022-2836(86)90385-2.

Taylor, S.C., 1995. Biochemical Process. US4508822A..

Wissner, J.L., Escobedo-Hinojosa, W., Heinemann, P.M., Hunold, A., Hauer, B., 2020. Methods for the detection and analysis of dioxygenase catalyzed dihydroxylation in mutant derived libraries. Enzyme Engineering and Evolution, Including Directed Evolution. Elsevier Inc., pp. 1–31. https://doi.org/10.1016/bs.mie.2020.04.022

Yanisch-Perron, C., Vieira, J., Messing, J., 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mpl8 and pUC19 vectors. Gene 33, 103–119. https://doi.org/10.1016/0378-1119(85)90120-9.

Yeh, W.K., Gibson, D.T., Liu, T.-N., 1977. Toluene dioxygenase: a multicomponent enzyme system. Biochem. Biophys. Res. Commun. 78, 401–410. https://doi.org/10.1016/0006-291X(77)91268-2.

Zylstra, G.J., Gibson, D.T., 1989. Toluene degradation by Pseudomonas putida F1. Nucleotide sequence of the todC1C2BADE genes and their expression in Escherichia coli. J. Biol. Chem. 264, 14940–14946.

Zylstra, G.J., McCombie, W.R., Gibson, D.T., Finette, B.A., 1988. Toluene degradation by Pseudomonas putida F1: genetic organization of the tod operon. Appl. Environ. Microbiol. 54, 1498–1503. https://doi.org/10.1128/aem.54.6.1498-1503.1988.

6.2 RESEARCH ARTICLE I

Supplementary Information

Supplementary Data

An enhanced toluene dioxygenase platform for the production of *cis*-1,2-dihydrocatechol in *Escherichia coli* BW25113 lacking glycerol dehydrogenase activity

Julian L. Wissner[a], Julian Ludwig[a], Wendy Escobedo-Hinojosa[a], Bernhard Hauer[a]*

[a]Institute of Technical Biochemistry, University of Stuttgart, Allmandring 31, 70569 Stuttgart, Germany

*Corresponding author

E-mail address: bernhard.hauer@itb.uni-stuttgart.de

1

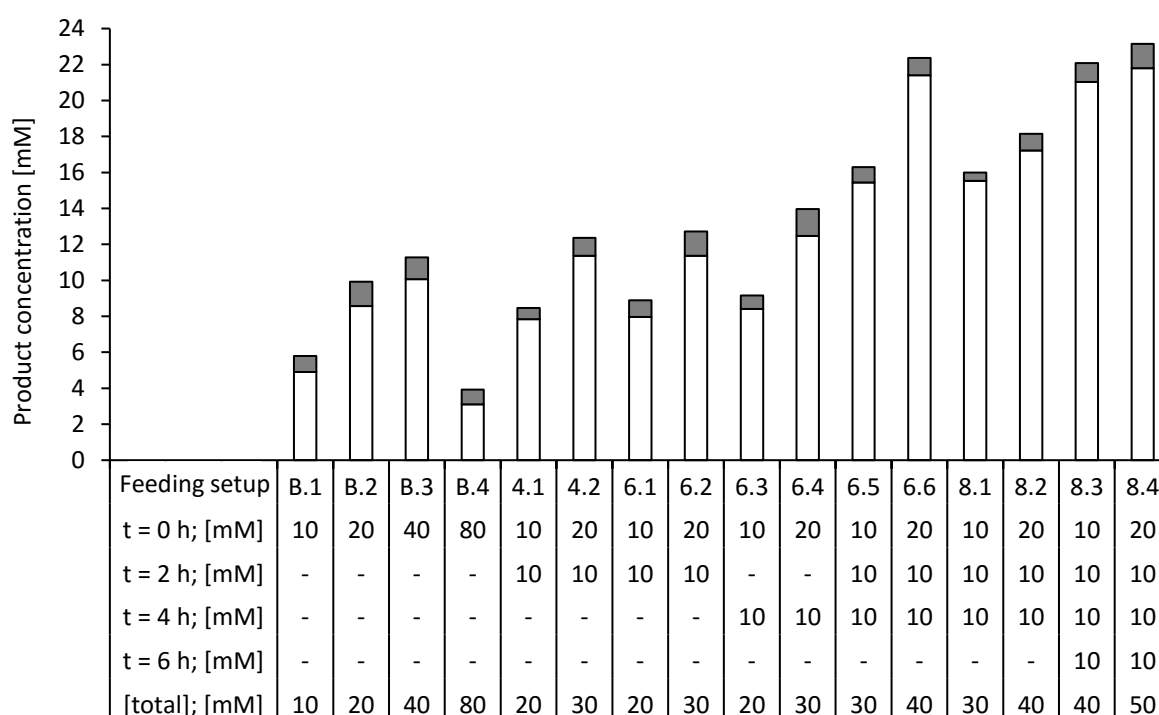# Supplementary section 1: Influence of substrate feeding on DHC formation

Feeding setup	B.1	B.2	B.3	B.4	4.1	4.2	6.1	6.2	6.3	6.4	6.5	6.6	8.1	8.2	8.3	8.4
t = 0 h; [mM]	10	20	40	80	10	20	10	20	10	20	10	20	10	20	10	20
t = 2 h; [mM]	-	-	-	-	10	10	10	10	-	-	10	10	10	10	10	10
t = 4 h; [mM]	-	-	-	-	-	-	-	-	10	10	10	10	10	10	10	10
t = 6 h; [mM]	-	-	-	-	-	-	-	-	-	-	-	-	-	-	10	10
[total]; [mM]	10	20	40	80	20	30	20	30	20	30	30	40	30	40	40	50

Fig. S1: Effect of different benzene starting concentrations and benzene-feeding strategies on the DHC formation. DHC is depicted in white, catechol in grey. TDO was expressed at 20°C for 20 h. Biotransformations were performed with 0.2 g_{cww} mL^{-1} cells in 0.1 M potassium phosphate buffer (pH 7.4) supplemented with 20 mM glucose and 10, 20, 40 or 80 mM starting benzene and the successive feeding of 10 mM benzene at different time points. The reaction was performed at 25°C and 180 rpm for 4, 6 or 8 h. B.1-B.4 = 4 h batch reaction, no benzene feeding (B.1 = 10 mM benzene, B.2 = 20 mM benzene; B.3 = 40 mM benzene, B.4 = 80 mM benzene); 4.1-4.2 = 4 h reaction, feed after 2 h (4.1 = 10 mM benzene starting concentration, 10 mM feed; 4.2 = 20 mM benzene starting concentration, 10 mM feed); 6.1-6.2 = 6 h reaction, feed after 2 h (6.1 = 10 mM benzene starting concentration, 10 mM feed; 6.2 = 20 mM benzene starting concentration, 10 mM feed); 6.3-6.4 = 6 h reaction, feed after 4 h (6.3 = 10 mM benzene starting concentration, 10 mM feed; 6.4 = 20 mM benzene starting concentration, 10 mM feed); 6.5-6.6 = 6 h reaction, feed after 2 h and 4 h (6.5 = 10 mM benzene starting concentration, two times 10 mM feed; 6.6 = 20 mM benzene starting concentration, two times 10 mM feed); 8.1-8.2 = 8 h reaction, feed after 2 h and 4 h (8.1 = 10 mM benzene starting concentration, two times 10 mM feed; 8.2 = 20 mM benzene starting concentration, two times 10 mM feed); 8.3-8.4 = 8 h reaction, feed after 2 h, 4 h and 6 h (8.3 = 10 mM benzene starting concentration, three times 10 mM feed; 8.4 = 20 mM benzene starting concentration, three times 10 mM feed).

2

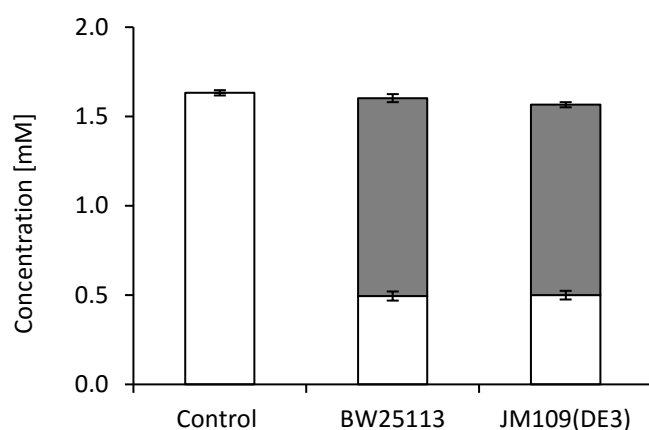# Supplementary section 2: Investigations on DHC dehydrogenation in *E. coli*

Fig. S2: Influence of two different *E. coli* strains regarding DHC degradation. DHC is depicted in white, catechol in grey. Control = reaction buffer containing 1.6 mM of DHC; BW25113 = non utilizing arabinose *E. coli* strain BW25113; JM109(DE3) = arabinose catabolizing *E. coli* strain JM109(DE3). Cells were grown over night at 37°C. Biotransformation were performed with 0.2 g_{cww} mL^{-1} cells in 0.1 M potassium phosphate buffer (pH 7.4) supplemented with 20 mM Glucose and 1.6 mM DHC at 30°C and 180 rpm for 1 h. Each bar is showing the mean value of three biological triplicates and the corresponding standard deviations.
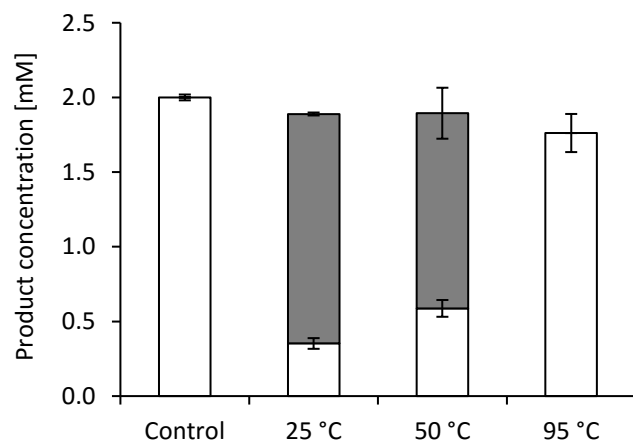

Fig. S3: Influence of different denaturation temperatures of BW25113 cell lysate for DHC conversion. DHC is depicted in white, catechol in grey. Control = reaction buffer containing 2.0 mM of DHC; 25°C = non-denaturized cell lysate; 50°C = cell lysate denaturized at 50°C for 10 min; 95°C = cell lysate denaturized at 95°C for 10 min. Cells were grown over night at 37°C. Biotransformation were performed with cell lysate of 0.2 g_{cww} mL^{-1} cells in 0.1 M potassium phosphate buffer (pH 7.4) supplemented with 20 mM glucose and 2.0 mM DHC at 30°C and 180 rpm for 20 h. Each bar is showing the mean value of three biological triplicates and the corresponding standard deviations.
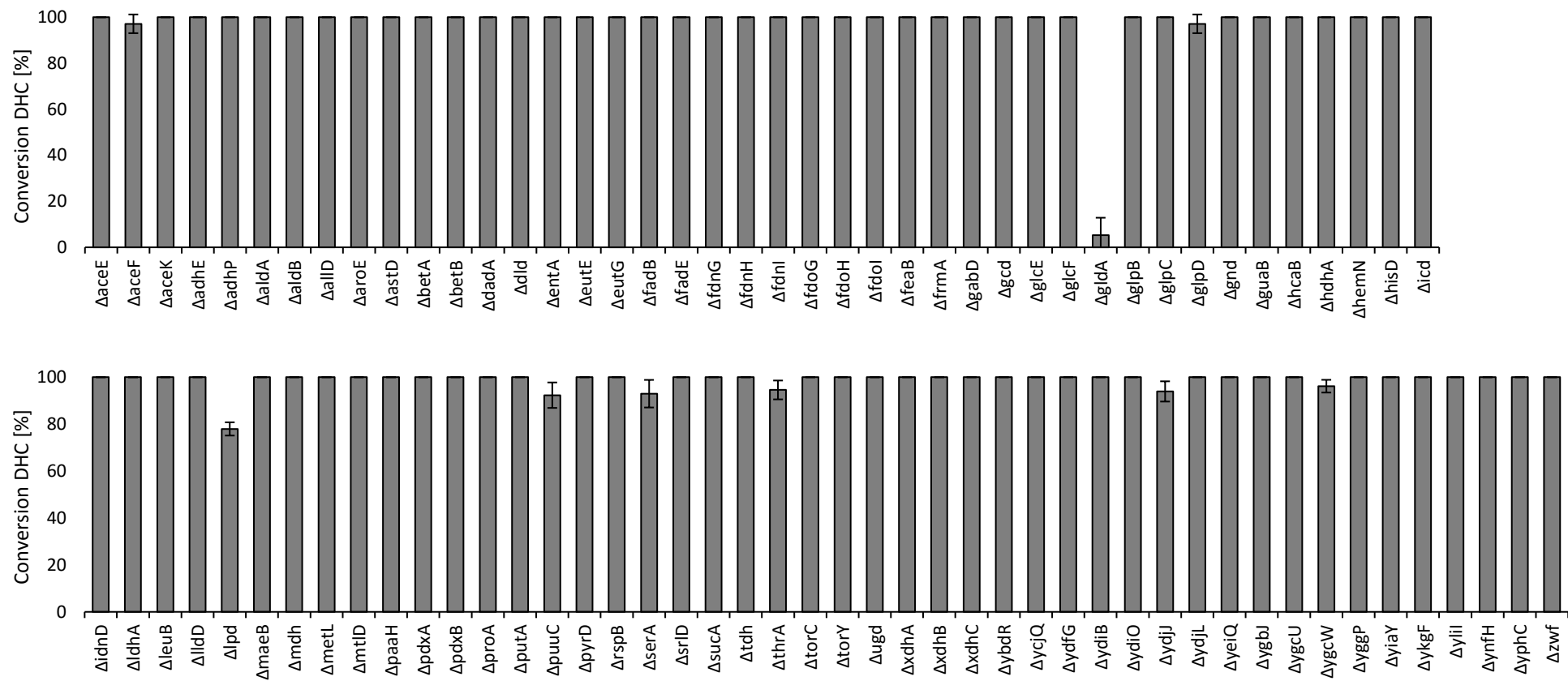
3

Fig. S4: Influence of different dehydrogenase deficient BW25113 strains on DHC degradation. DHC conversion is shown in grey. The deleted dehydrogenase genes are sorted alphabetically. Cells were grown for 20 h at 37°C in LB medium, before cultivated in TB medium at 37°C for 20 h. Biotransformations were performed in 0.1 M potassium phosphate buffer (pH 7.4) supplemented with 20 mM glucose and 2.0 mM DHC at 30°C and 800 rpm for 20 h in 96-deep well plates. Each bar is showing the mean value of three biological triplicates and the corresponding standard deviations.

Table S1: Catalog of all tested dehydrogenase deficient KEIO knock out strains.

ECK number	Knocked out gene	JW id	*E. coli* b number
ECK0113	aceE	JW0110	b0114
ECK0114	aceF	JW0111	b0115
ECK4008	aceK	JW3976	b4016
ECK1235	adhE	JW1228	b1241
ECK1472	adhP	JW1474	b1478
ECK1408	aldA	JW1412	b1415
ECK3577	aldB	JW3561	b3588
ECK0510	allD	JW0505	b0517
ECK3268	aroE	JW3242	b3281
ECK1744	astD	JW5282	b1746
ECK0309	betA	JW0303	b0311
ECK0310	betB	JW0304	b0312
ECK1177	dadA	JW1178	b1189
ECK2126	dld	JW2121	b2133
ECK0589	entA	JW0588	b0596
ECK2450	eutE	JW2439	b2455
ECK2448	eutG	JW2437	b2453
ECK3838	fadB	JW3822	b3846
ECK0222	fadE	JW5020	b0221
ECK1468	fdnG	JW1470	b1474
ECK1469	fdnH	JW1471	b1475
ECK1470	fdnI	JW1472	b1476
ECK3887	fdoG	JW3865	b3894
ECK3886	fdoH	JW3864	b3893
ECK3885	fdoI	JW3863	b3892
ECK1382	feaB	JW1380	b1385
ECK0353	frmA	JW0347	b0356
ECK2655	gabD	JW2636	b2661
ECK0123	gcd	JW0120	b0124
ECK2973	glcE	JW5487	b4468
ECK2972	glcF	JW5486	b4467
ECK3937	gldA	JW5556	b3945
ECK2234	glpB	JW2236	b2242
ECK2235	glpC	JW2237	b2243
ECK3412	glpD	JW3389	b3426
ECK2024	gnd	JW2011	b2029
ECK2504	guaB	JW5401	b2508

ECK2538	hcaB	JW2525	b2541
ECK1614	hdhA	JW1611	b1619
ECK3860	hemN	JW3838	b3867
ECK2015	hisD	JW2002	b2020
ECK1122	icd	JW1122	b1136
ECK4260	idnD	JW4224	b4267
ECK1377	ldhA	JW1375	b1380
ECK0075	leuB	JW5807	b0073
ECK3595	lldD	JW3580	b3605
ECK0115	lpd	JW0112	b0116
ECK2458	maeB	JW2447	b2463
ECK3225	mdh	JW3205	b3236
ECK3932	metL	JW3911	b3940
ECK3589	mtlD	JW3574	b3600
ECK1392	paaH	JW1390	b1395
ECK0053	pdxA	JW0051	b0052
ECK2314	pdxB	JW2317	b2320
ECK0244	proA	JW0233	b0243
ECK1005	putA	JW0999	b1014
ECK1295	puuC	JW1293	b1300
ECK0936	pyrD	JW0928	b0945
ECK1575	rspB	JW1572	b1580
ECK2909	serA	JW2880	b2913
ECK2700	srlD	JW2674	b2705
ECK0714	sucA	JW0715	b0726
ECK3606	tdh	JW3591	b3616
ECK0002	thrA	JW0001	b0002
ECK0987	torC	JW0981	b0996
ECK1874	torY	JW1862	b1873
ECK2023	ugd	JW2010	b2028
ECK2862	xdhA	JW5462	b2866
ECK2863	xdhB	JW2835	b2867
ECK2864	xdhC	JW2836	b2868
ECK0602	ybdR	JW0601	b0608
ECK1308	ycjQ	JW1306	b1313
ECK1532	ydfG	JW1532	b1539
ECK1690	ydiB	JW1682	b1692
ECK1693	ydiO	JW5275	b1695
ECK1772	ydjJ	JW1763	b1774

6

ECK1774	ydjL	JW1765	b1776
ECK2166	yeiQ	JW2160	b2172
ECK2731	ygbJ	JW2706	b2736
ECK2767	ygcU	JW5442	b4463
ECK2768	ygcW	JW5443	b2774
ECK2927	yggP	JW5477	b4465
ECK3578	yiaY	JW5648	b3589
ECK0306	ykgF	JW0300	b0307
ECK0827	yliI	JW0821	b0837
ECK1585	ynfH	JW5261	b1590
ECK2542	yphC	JW5842	b2545
ECK1853	zwf	JW1841	b1852

Table S2: Gene names and references. Link provided is aiming to address the reader to the specific strain and the corresponding protein cited on the table.

Gene name	Strain	Access number	NCBI Link
todD	*P. putida* F1	GenBank AAA26009.1	https://www.ncbi.nlm.nih.gov/protein/AAA26009
entA	[a]*E. coli* K-12	NCBI Reference Sequence: NP_415128.1	https://www.ncbi.nlm.nih.gov/protein/NP_415128.1
hcaB	[a]*E. coli* K-12	NCBI Reference Sequence: NP_417036.1	https://www.ncbi.nlm.nih.gov/protein/NP_417036.1
bedD	*P. putida* ML2	GenBank: AAC44426.1	https://www.ncbi.nlm.nih.gov/protein/AAC44426.1
gldA	[a]*E. coli* K-12	NCBI Reference Sequence: NP_418380.4	https://www.ncbi.nlm.nih.gov/protein/NP_418380.4

a; K-12 is the direct parent strain of BW25113

Table S3: Substrate spectrum of GldA and BedD.

Protein	Natural host	Substrate scope	Reference
GldA (glycerol dehydrogenase)	[a]*E. coli* K-12	glycerol, D-1-Amino-2-propanol, 1,2-Propanediol, 2,3-butanediol, 1,2-butanediol, 3-amino-1,2-propanediol, 1,2-ethanediol	Campbell et al., (1978); Kelley and Dekker, (1985); Kelley and Dekkers, (1984)
BedD (benzene dehydrogenase)	*P. putida* ML2	glycerol, 1,2-propandiol, *cis*-1,2-dihydrocatechol, *cis*-1,2-cyclohexanediol, *cis*-toluene dihydrodiol, *cis*-naphthalene dihydrodiol	Fong et al., (1996); Fong and Tan, (1999)

a; K-12 is the direct parent strain of BW25113

Table S4: Homology table. Identity was obtained by alignment of the proteins on UniProt.

	TodD (*P. putida*)	EntA ([b]*E. coli*)	HcaB ([b]*E. coli*)	BedD (*P. putida*)	GldA ([b]*E. coli*)	
TodD (*P. putida* F1)	100	23	43	17	15	
EntA ([b]*E. coli* K-12)		100	21	13	13	Identity
HcaB ([b]*E. coli* K-12)			100	10	12	
[a]BedD (*P. putida* ML2)				100	73	
GldA ([b]*E. coli* K-12)					100	

a; in literature also known as glyceroldehydrogenase GldA

b; K-12 is the direct parent strain of BW25113

Supplementary section 3: Assessment of the presence and the lack of *gldA* gene in *E. coli* BW25113 wild type and BW25113 Δ*gldA*

E. coli competent cells of strains BW25113 wild type and BW25113 Δ*gldA* were regenerated for 30 min at room temperature in LB media, plated out on LB agar plates (without antibiotics) and incubated over night at 37°C. A 5 mL LB preculture was inoculated with a single colony from each plate, incubated for 2 h at 37°C and 180 rpm. Afterwards, the total volume was used to inoculate a 250 mL shaking flask containing 50 mL TB media and incubated over night at 37°C and 180 rpm. The cells were harvested at 4°C at 4000 x *g* for 20 min. Genomic DNA purification was performed according to the Kit instructions *Pure Link, Genomic DNA Mini Kit* (Invitrogen, Carlsbad, US). The isolated gDNA was diluted to a concentration of 80 ng μL^{-1} and used as template following the PCR recipe shown in Table 6S, with primers listed in table 5S, using KOD hot start DNA polymerase (Novagen, Madison, US). The PCR amplicons were generated by using the PCR program shown in table 7S.

Table S5: Primer list; if = intern forward, ir = intern reverse, ef = extern forward, er = extern reverse.

No.	Primer	Sequence
I	jwi_gDNA_gldA_if	ATGGACCGCATTATTCAATCACCGGG
II	jwi_gDNA_gldA_ir	TTATTCCCACTCTTGCAGGAAACGC
III	jwi_gDNA_gldA_ef	CGGTTCAGGAGCTGCAAACGC
IV	jwi_gDNA_gldA_er	TCACAGATTCGACCTTCCGGGC

Table S6: PCR mix

PCR mix	[μL]
10 x KOD buffer	5.0
25 mM magnesium sulfate	4.5
dNTP mix (2 mM each)	5.0
Forward primer	1.0
Reverse primer	1.0
Template	1.0
KOD	1.0
DMSO	2.5
ddH$_2$O	29.0

Table S7: PCR programm.

No.	Name	T [°C]	t [min]
1	Initial denaturation	95	5
2	Denaturation	95	0.5
3	Annealing	60	0.5
4	Extention	70	1.5
5	Repetition with 35 cycles, go to no. 2	-	-
6	Final extention	70	5
7	Hold	8	-

For the identification of the PCR products an electrophoretic analysis was performed (1.0% agarose gel). The agarose gel was loaded with 5 µL ddH$_2$O, 2.5 µL 6 x DNA loading buffer and 10 µL of the PCR product. The gel was run at 100 V for 1 h (Fig. S5).
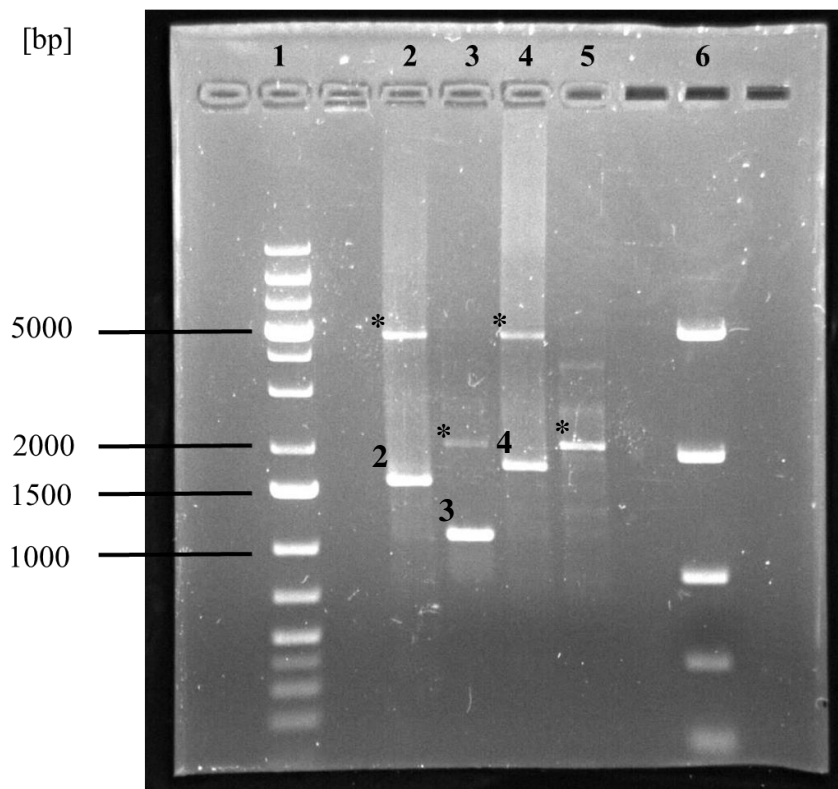
Fig. S5: Agarose gel for assessment of the presence and the lack of *gldA* gene. Bands 2 and 3 were obtained with the wild type gDNA, bands 4 and 5 with the Δ*gldA* gDNA. Products 2 and 4 were amplified with primers III and IV, products 3 and 5 with primers I and II. 1 = gene ruler 1 kb plus; 2 = a 1594 bp amplicon containing the *gldA* gene (1104 bp) plus upstream and dowstream regions (240 and 250 bp, respectively); 3 = only the gldA gene (1104 bp); 4 = amplicon containing the kanamycin cassette replacing the *gldA* gene in the mutant strain, plus upstream and dowstream regions; 5 = no 1104 bp PCR product, indicating effective delition of the *gldA* gene; 6 = Fast ruler DNA ladder. Unspecific products are indicated with asterisks.

Supplementary section 4: Investigations on TDO toluene conversion and *cis*-1,2-dihydro-3-methylcatechol dehydration in *E. coli*

HPLC-DAD set up for 3-methylcatechol

For toluene biotransformations, with MTBE extracted samples were analysed by high-pressure liquid chromatography (HPLC). An Agilent 1260 Infinity II, Santa Clara, US) equipped with a C18 column (Supelco C18 Discovery, 5 μm, 4.0 x 150 mm, Bellefonte, US) and a diode array detector (Agilent 1260 Infinity II DAD HS, Santa Clara, US) was operated isothermally at 30°C. Measurements were run at a flow rate of 1.0 mL min^{-1}. The mobile phase was water/acetonitrile with a linear gradient of t = 0 min, 90/10 (v/v); t = 1.0 min, 90/10 (v/v); t = 3.0 min, 20/80 (v/v); t = 6.0 min, 20/80 (v/v); t = 6.01 min, 90/10 (v/v); t = 10.0 min, 90/10 (v/v). Peak areas were measured by the integrator and transformed into concentration using the standard curve. A potassium phosphate buffer solution (pH 7.4), containing the pure standard; 3-methylcatechol was extracted with MTBE to obtain the samples for the standard curve. The employed wavelength to detect *cis*-1,2-dihydro-3-methylcatechol and 3-methylcatechol were 262 and 210 nm, respectively.
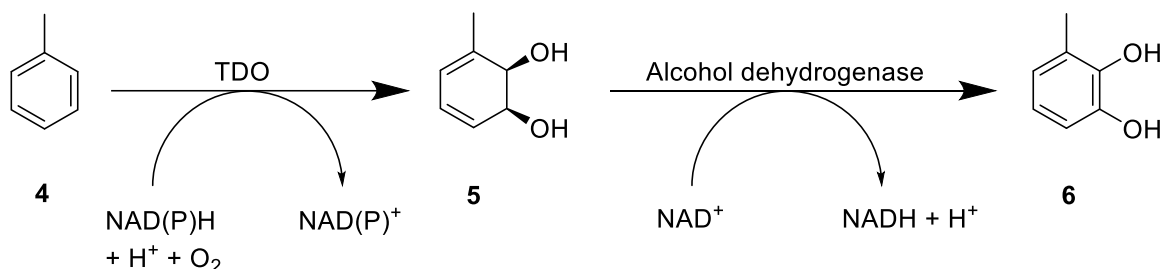
Fig. S6: Toluene conversion by TDO from *P. putida* F1 and the subsequent downstream reaction driven by a native alcohol dehydrogenase from *E. coli* BW25113 observed in this work. Dearomatizing dihydroxylation of toluene **4**, generating *cis*-1,2-dihydro-3-methylcatechol **5** and the dehydrogenation reaction, resulting in 3-methylcatechol **6** formation.
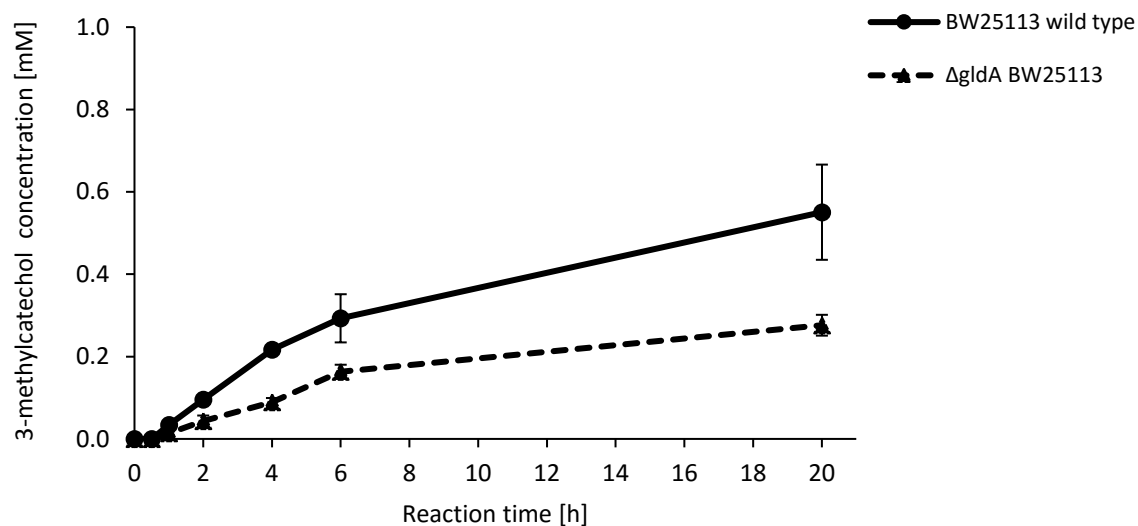
Fig. S7: 3-Methylcatechol production in BW25113 and BW25113 Δ*gldA* during toluene biotransformation. Wild type *E. coli* BW25113 is depicted as solid line, BW25113 Δ*gldA* as dotted line. Protein was expressed at 20°C for 20 h. Biotransformations were performed with 0.2 g_{cww} mL^{-1} cells in 0.1 M potassium phosphate buffer (pH 7.4) supplemented with 20 mM Glucose and 10 mM toluene at 30°C and 180 rpm. Each bar is showing the mean value of three biological triplicates and the corresponding standard deviations.
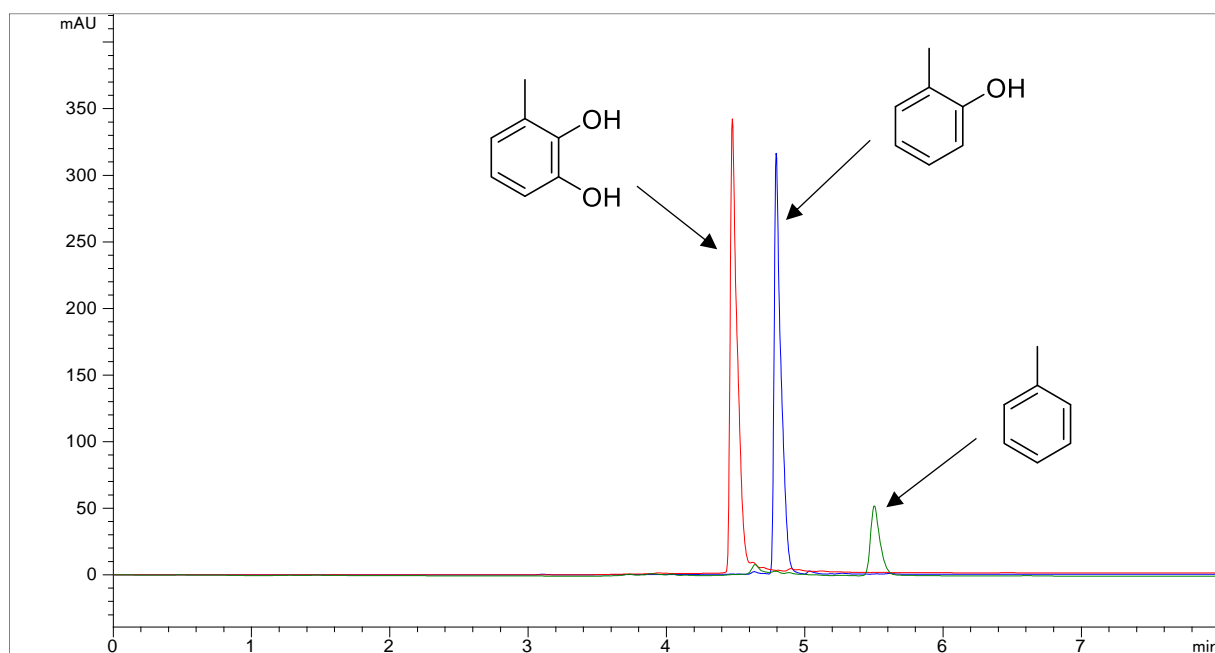
Fig. S8: HPLC-DAD (262 nm) chromatogram of commercial standards for toluene biotransformation. 3-methylcatechol (4.51 min; red), *o*-cresol (4.83 min; blue) and toluene (5.55 min; green).

12
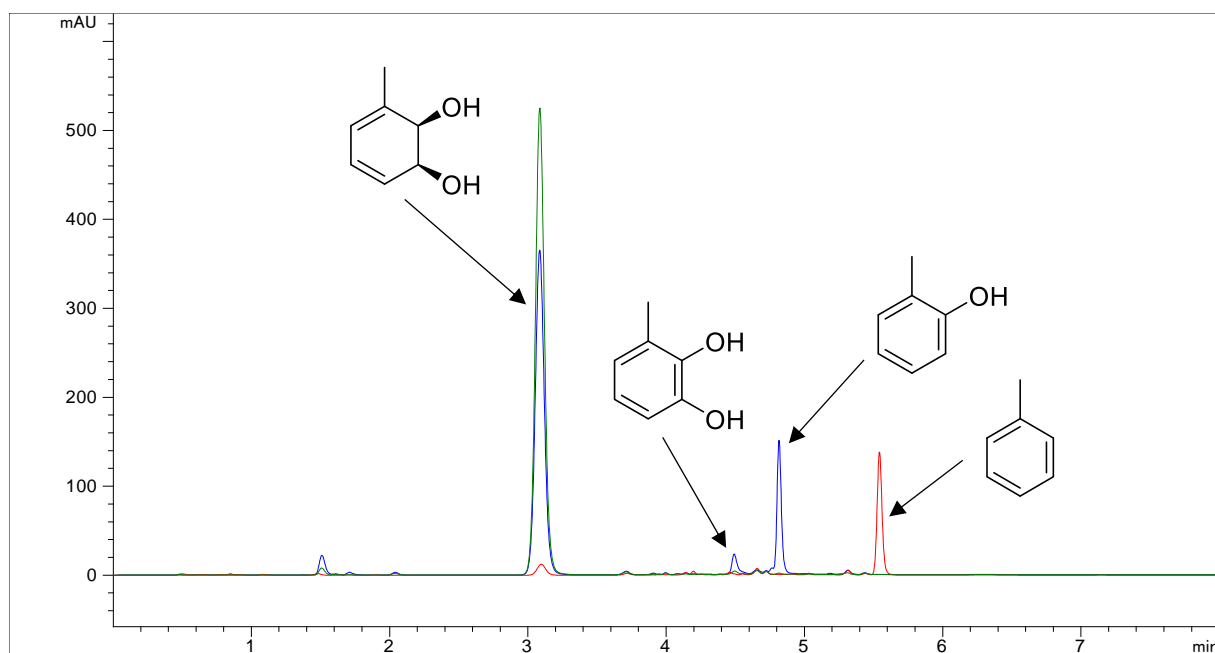
Fig. S9: HPLC-DAD (262 nm) chromatogram of the toluene biotransformation after 5 min (red), 2 h (green) and 20 h (blue). *Cis*-1,2-dihydro-3-methylcatechol (3.08 min), 3-methylcatechol (4.51 min), *o*-cresol (4.83 min) and toluene (5.55 min). Above a reaction time of 6 h, *cis*-1,2-dihydro-3-methylcatechol was degraded spontaneously to *o*-cresol.

13

Section 5: HPLC-DAD chromatograms of *cis*-1,2-dihydrocatechol and catechol
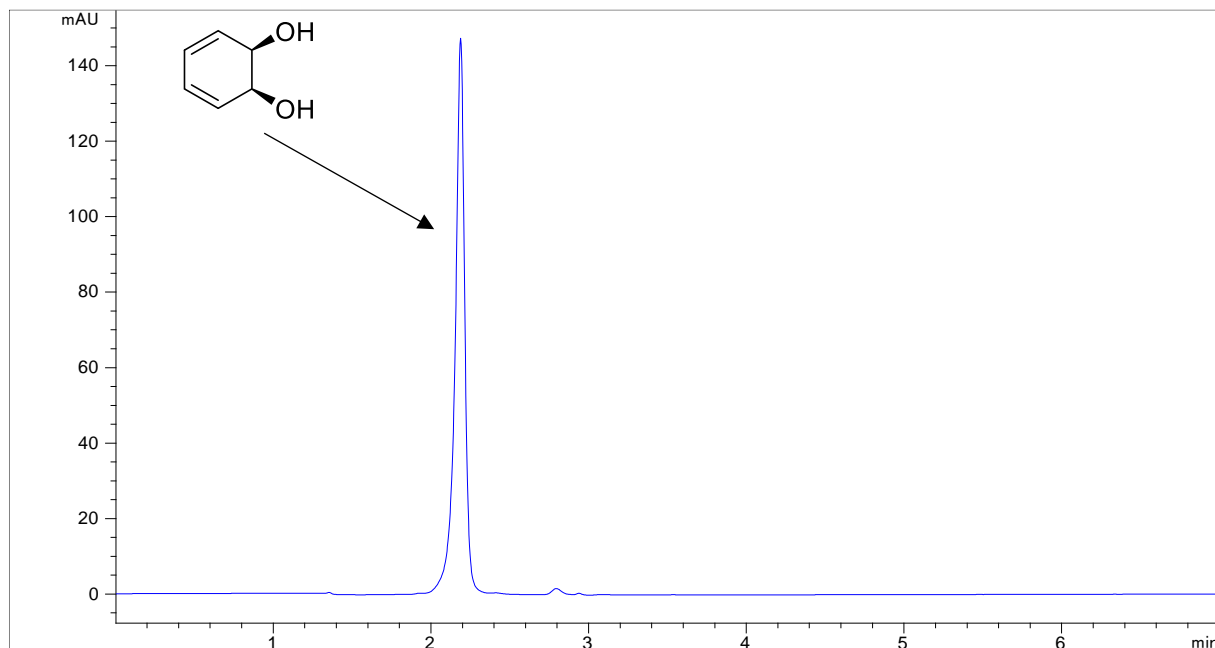
Fig. S10: HPLC-DAD (262 nm) chromatogram of DHC standard (2.19 min), synthesized *via* semi-preparative biotransformation.
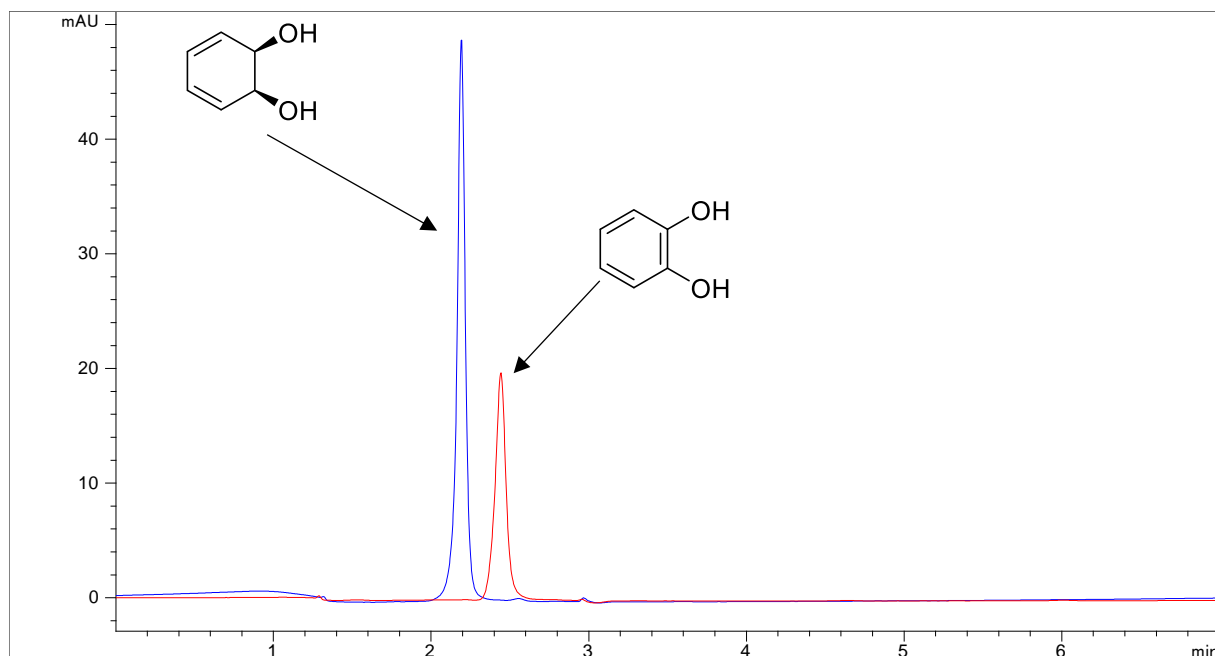
Fig. S11: HPLC-DAD (262 nm) chromatogram of DHC standard (2.19 min; blue) and catechol standard (2.44 min; red).
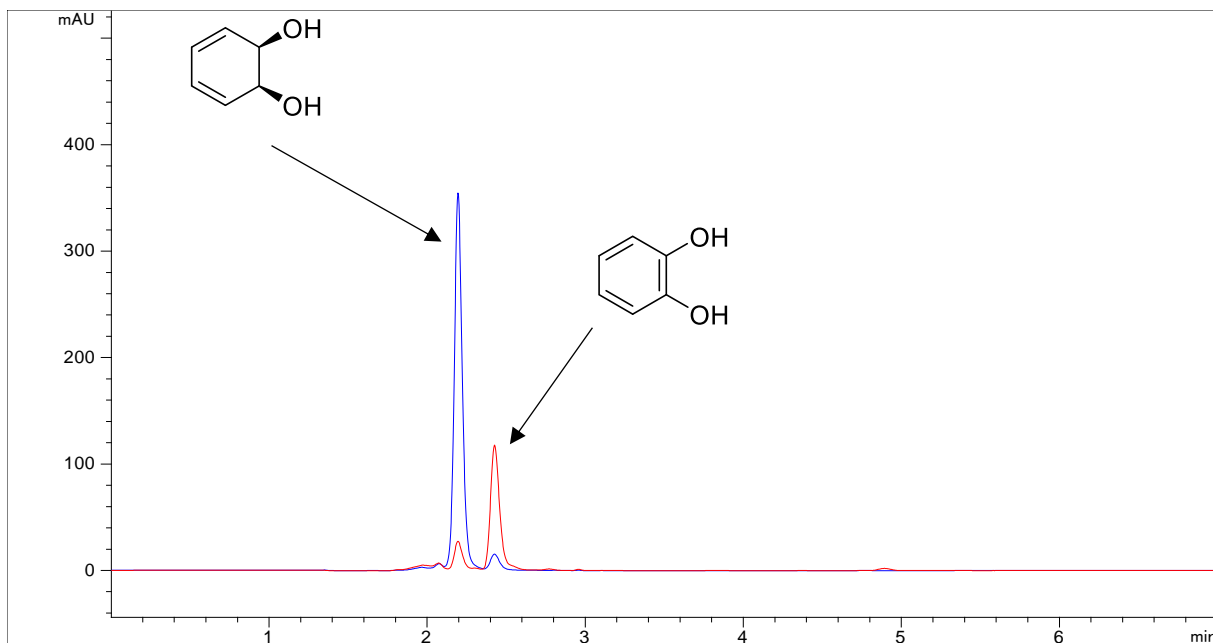
Fig. S12: HPLC-DAD (262 nm) chromatogram of standard biotransformation after 4 h (red) and optimized biotransformation (20°C expression and 25°C biotransformation) after 4 h (blue).
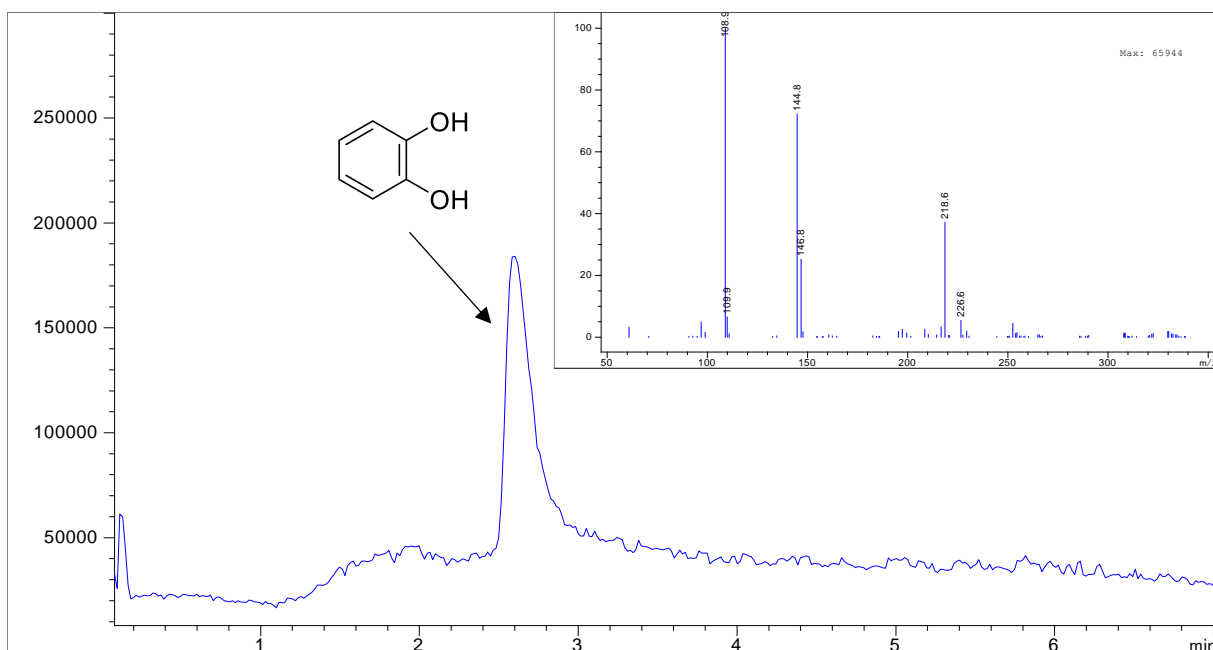
Fig. S13: HPLC-ESI-MS chromatoram of catechol standard (2.58 min), measured in negative scan. Fragmentation pattern in right corner (109 [M-H]; 145 [M+Cl]; 219 [2M-H]).
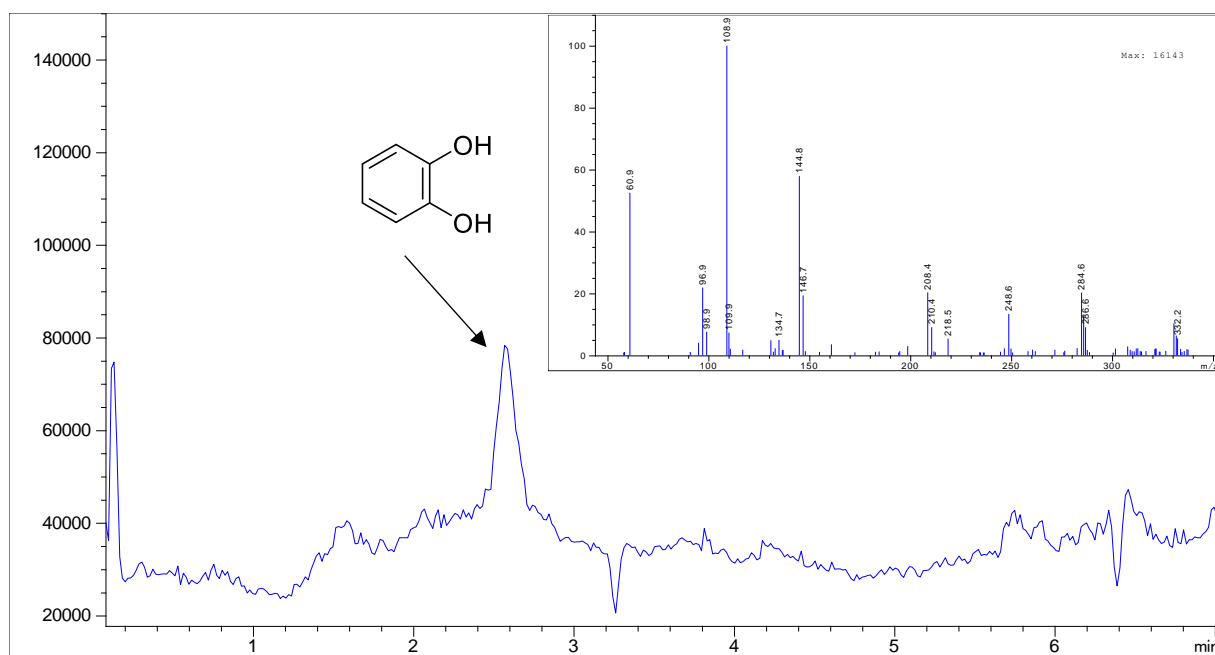
Fig. S14: HPLC-ESI-MS chromatogram of biotransformation after 4 h measured in negative scan. Fragmentation pattern in right corner (109 [M-H]; 145 [M+Cl]; 219 [2M-H]).

Section 6: NMR; spectroscopic analysis
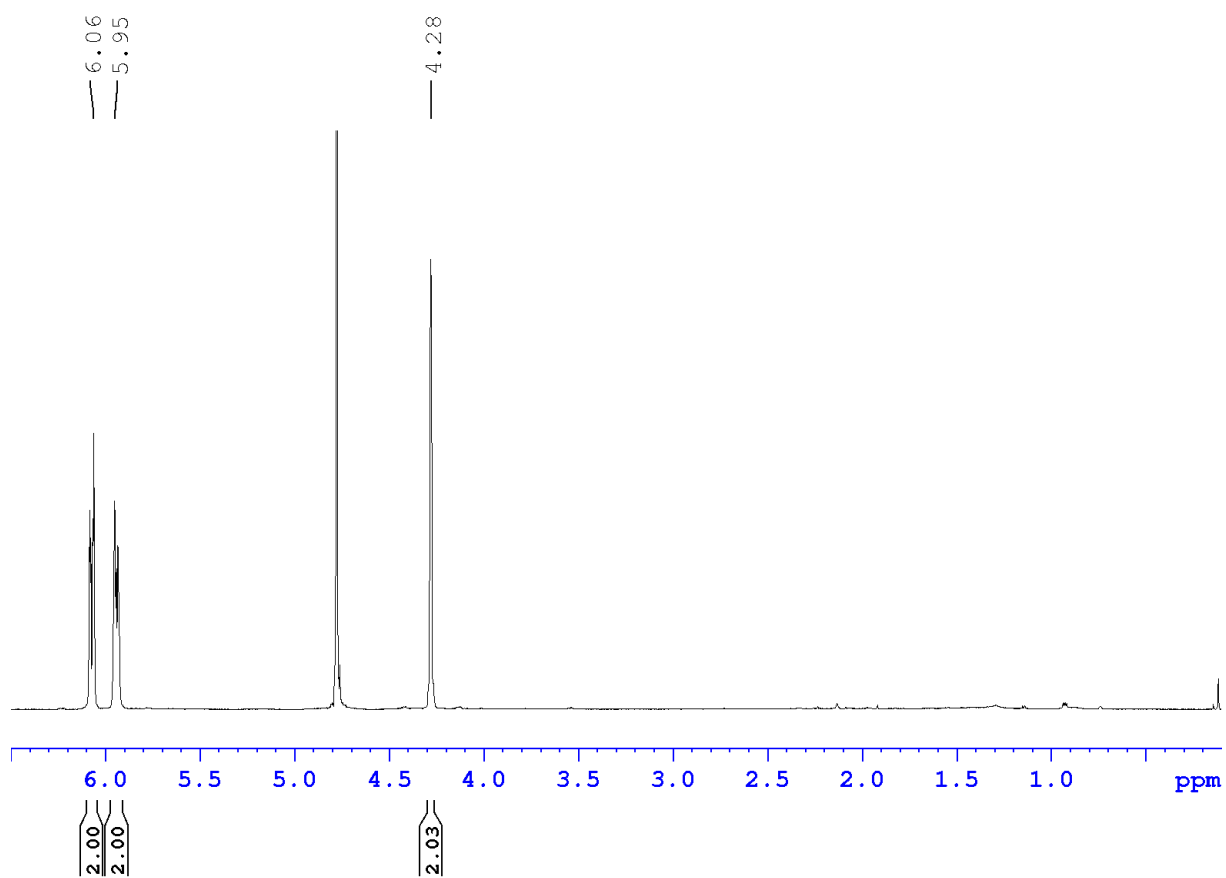
Fig. S15: ^{1}H-NMR of DHC in D$_2$O. TMS was used as standard. H$_2$O can be seen at 4.79 ppm.
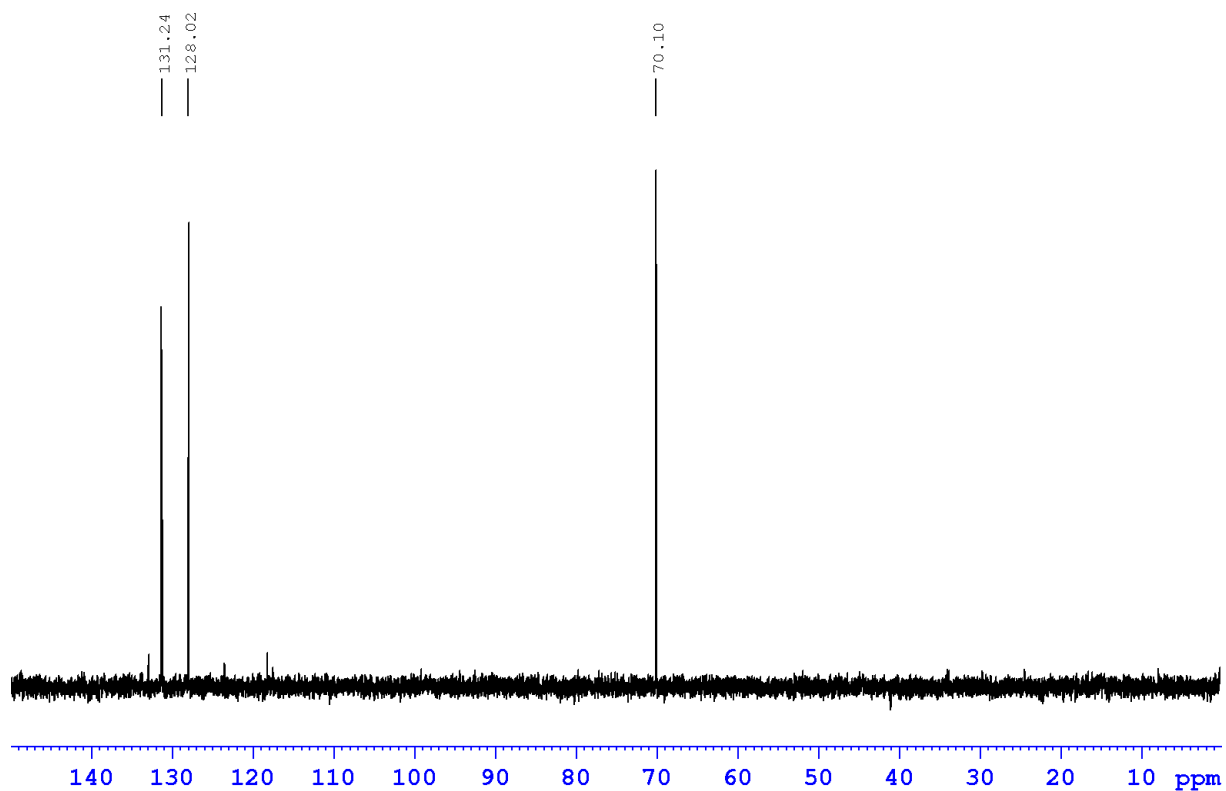
Fig. S16: ^{13}C-NMR of DHC in D$_2$O. TMS was used as standard.

17

References

Campbell, R.L., Swain, R.R., Dekker, E.E., 1978. Purification, separation, and characterization of two molecular forms of D-1-amino-2-propanol:NAD+ oxidoreductase activity from extracts of Escherichia coli K-12. J. Biol. Chem. 253, 7282–7288.

Fong, K.P.Y., Goh, C.B.H., Tan, H.M., 1996. Characterization and expression of the plasmid-borne bedD gene from Pseudomonas putida ML2, which codes for a NAD+-dependent cis-benzene dihydrodiol dehydrogenase. J. Bacteriol. 178, 5592–5601. https://doi.org/10.1128/jb.178.19.5592-5601.1996

Fong, K.P.Y., Tan, H.M., 1999. Characterization of a novel cis-benzene dihydrodiol dehydrogenase from Pseudomonas putida ML2. FEBS Lett. 451, 5–9. https://doi.org/10.1016/S0014-5793(99)00520-7

Kelley, J.J., Dekker, E.E., 1985. Identity of Escherichia coli D-1-amino-2-propanol:NAD+ oxidoreductase with E. coli glycerol dehydrogenase but not with Neisseria gonorrhoeae 1,2-propanediol:NAD+ oxidoreductase. J. Bacteriol. 162, 170–175. https://doi.org/10.1128/jb.162.1.170-175.1985

Kelley, J.J., Dekkers, E.E., 1984. D- l-Amino-2-propanol : NAD + Oxidoreductase. Society 259, 2124–2129.

6.3 RESEARCH ARTICLE II

Manuscript

Contents lists available at ScienceDirect

Journal of Biotechnology

An engineered toluene dioxygenase for a single step biocatalytical production of (-)-(1*S*,2*R*)-*cis*-1,2-dihydro-1,2-naphthalenediol

Julian L. Wissner [a], Wendy Escobedo-Hinojosa [a], Andreas Vogel [b], Bernhard Hauer [a],*

[a] *Institute of Technical Biochemistry, University of Stuttgart, Allmandring 31, 70569, Stuttgart, Germany*
[b] *c-LEcta GmbH, Perlickstr. 5, 04103, Leipzig, Germany*

ARTICLE INFO

ABSTRACT

cis-1,2-Dihydro-1,2-naphthalenediol (DHND) is a valuable molecule employed for the pharmaceutical synthesis of bioactive compounds, such as bicyclic conduritol analogues. Enantiopure (+)-(1*R*,2*S*)-DHND (>98 % ee) is easily biosynthesized through the dearomatizing dihydroxylation of naphthalene, catalyzed by toluene dioxygenase (TDO) from *Pseudomonas putida* F1. However, the opposite enantiomer (-)-(1*S*,2*R*)-DHND could not be directly accessed, neither by chemical synthesis nor *via* biocatalytic approaches. Herein, we report a one-step biosynthesis of the opposite enantiomer (-)-(1*S*,2*R*)-DHND in a recombinant TDO *E. coli* BW25113 platform. We based on a semi-rational approach to generate a set of TDO variants, targeting exclusively the hotspot position F366, in order to enable an enantiomeric switch in the generated product. Eight out of nine single point variants were active and showed not only an alteration in enantioselectivity, but also generated an enantiomeric excess of the pursued product. Variant TDO_{F366V} outperformed above the rest of the set, enabling the synthesis of (-)-(1*S*,2*R*)-DHND not only with an excellent enantiomeric excess of 90 %, but also with an advantageous product formation. A comparative semi-preparative biosynthesis yielded, 287 mg of (+)-(1*R*,2*S*)-DHND (>98 % *ee*) and 101 mg of (-)-(1*S*,2*R*)-DHND (90 % *ee*), when performed in a total volume of 100 mL with TDO wild-type and TDO_{F366V} resting cells, respectively.

1. Introduction

Chiral *cis*-dihydrodiendiols of aromatic compounds are often used as synthons in organic chemistry for the production of fine chemicals, pharmaceuticals and bioactive compounds (Endoma et al., 2002; Gally et al., 2015; Hudlicky et al., 1989; Wissner et al., 2020). Especially the hydroxylated synthons of naphthalene and its derivatives, are relevant for the generation of compounds with therapeutic application (Boyd et al., 1996; Kwit et al., 2008; Sello and Orsini, 2005). The enantiomer (+)-(1*R*,2*S*) of *cis*-1,2-dihydro-1,2-naphthalenediol **2a** (Fig. 1A), is a useful scaffold for the synthesis of various natural products (Fig. 1B).

The chemical synthesis addressing the generation of **2a** consists in four chemical steps, using as initial substrate 2-vinylbenzaldehyde (Mukherjee et al., 2010). Disadvantages of the chemical synthesis are not only in terms of poor yield (29 %) and reduced enantiomeric excess, but also that the initial substrate is rather expensive. Besides, harsh reaction conditions and organic solvents are needed, as well as, the use of toxic heavy metal catalysts such as PdCl_2(MeCN)_2 and Ru-based Grubbs-II catalyst.

The biocatalytic alternative to generate enantiopure **2a** comprises a one-step reaction, utilizing the accessible molecule naphthalene as initial substrate and a Rieske non-heme iron dioxygenase (RO) as biocatalyst, under mild reaction conditions, employing water as solvent. In addition to the RO naphthalene dioxygenase (NDO) from *Pseudomonas sp.* NCIB 9816-4, toluene dioxygenase (TDO) from *P. putida* F1 is another well characterized RO system for the dearomatizing dihydroxylation of naphthalene. In a highly efficient one-step reaction, TDO converts, in excellent yield, naphthalene to the enantiopure product **2a** (> 98 % *ee*), (Endoma et al., 2002; McIver et al., 2008). Thus, generation of **2a** by the biocatalytic approach (one-step, > 98 % *ee*), surpasses the chemical synthesis (four-steps, 90 % *ee*), regarding the cost/benefit balance and in terms of efficiency, also in a more sustainable and environmental friendly manner.

Despite the numerous applications of synthon **2a**, in literature there is not a single application or reaction described for the opposite enantiomer **2b**. This is merely due to the commercial unavailability of enantiomer **2b**. Until date, there is no chemical or enzymatic reaction, yielding directly enantiopure **2b** in tangible amounts.

---

* Corresponding author.
 E-mail address: bernhard.hauer@itb.uni-stuttgart.de (B. Hauer).

A

TDO wild-type → **2a** — 100% product formation — >98% *ee* (+)-(1R,2S)

TDO$_{F366V}$ → **2b** — 81% product formation — 90% *ee* (-)-(1S,2R)

1

B
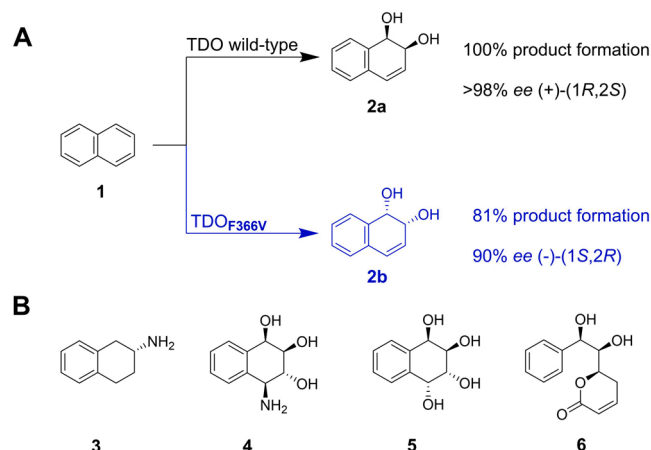
3 **4** **5** **6**

Fig. 1. A. (+)-(1R,2S)-1,2-dihydro-1,2-naphthalenediol **2a** (black) is generated from naphthalene **1** by TDO wild-type. The opposite enantiomer (-)-(1S,2R)-1,2-dihydro-1,2-naphthalenediol **2b** (blue) is generated, as well, from **1** by the engineered variant TDO$_{F366V}$. **B** Selected examples of pharmaceutically relevant synthetized compounds, employing **2a** as scaffold. The amphetamine substitute (R)-2-aminotetraline **3** (Orsini et al., 2002), the bicyclic conduritol analogues **4** and **5**, which their respective derivatives have antibiotic and anti-tumor activities (Orsini et al., 2004), and the cytotoxic compound (+)-goniodiol **6** (Banwell et al., 2003). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

A multistep approach combining organic synthesis and biocatalysis for the generation of **2b** have only achieved a reduced total product yield of only 19 %. For this route, a three step synthesis of racemic *cis*-1,2-dihydro-1,2-naphthalenediol (Jeffrey et al., 1974), is followed by a kinetic resolution performed either by *P. putida* 119 or *P. putida* 11767 (Allen et al., 1995; Jeffrey et al., 1975).

The aim of our work was to enable a semi-preparative, one-step, biosynthesis of **2b**, employing naphthalene as substrate and TDO as biocatalyst. In order to make feasible our goal, we generate a set of TDO variants targeting the hot-spot position F366, pursuing to identify the most suitable one for an efficient biocatalytic synthesis of **2b**.

2. Materials and methods

Description of materials, bacterial strain, plasmid, growth conditions, protein expression, variant generation, biotransformation analysis methods, as well as, identification and characterization of **2a** and **2b** are given in the supplementary information.

3. Results and discussion

Our experimental approach relied on previous ROs-based exploratory approaches demonstrating the viability to generate the enantiomer **2b**. Prior to this work it was reported, that by exchanging phenylalanine residue 352 in NDO for a valine, a mild shift in product enantioselectivity was achieved, enabling the generation of small amounts of **2b**, but favoring still the formation of product **2a** (84 % *ee* for **2a**), (Parales et al., 2000). This position showed to have a major influence in enantio-, as well as in regioselectivity not only for naphthalene, but also for the aromatic compounds biphenyl and phenanthrene. The shared homology between TDO and NDO led to transferring such findings to TDO by introducing the same mutation in the equivalent position 366 (Vila et al., 2017). TDO$_{F366V}$ lead to drastical changes in enantio- and regioselectivity. The enantioselectivity shift was observed when bromobenzene, styrene, and indene were tested as substrates. Mutation TDO$_{F366V}$ exerted the most dramatic influence when indene was used as substrate. Regarding the enantiomeric excess, a switch from +33 *cis*-(1S, 2R)-indandiol to 9 % for *cis*-(1R,2S)-indandiol was achieved. Though,

alterations regarding chemo- and enantioselectivity were observed in a substrate dependent manner, naphthalene was not included in the substrate panel. Besides, only TDO$_{F366V}$ was addressed and left unexplored all the rest of possible amino acid exchanges at position 366. In this work, we took such previous knowledge as our foundation to generate a set of variants targeting the hot-spot position 366 of TDO. We substituted the TDO 366 phenylalanine residue with alanine, isoleucine, leucine, asparagine, glutamine, serine, threonine, valine, and tyrosine.

All tested variants except TDO$_{F366Y}$, were active towards naphthalene and enabled not only a shift in enantioselectivity, but also generated the desired enantiomer **2b** as main product (Table 1).

Specially, TDO$_{F366V}$ was of great interest, since it not only showed an excellent enantiomeric excess of 90.1 % for **2b**, but also a very good product formation of 80.7 % (8.07 mM) (Table 1). Our findings are in agreement with respect to the reported behavior observed in NDO F352 variants (Parales et al., 2000). For instance, in both enzymes the exchange into a leucine at equivalent positions, resulted in the highest activity (96 % for NDO, 99 % for TDO), followed in both cases by the exchange into a valine (83 % for NDO, 81 % for TDO). The lowest activity was observed for the alanine substitution (49 % for NDO, 2 % for TDO). In addition, the most striking enantioselectivity changes were performed by the variants harboring a valine instead of a phenylalanine (84 % *ee* for **2a** with NDO, 90 % *ee* for **2b** in TDO), followed by the threonine exchange (86 % *ee* for **2a** with NDO, 84 % *ee* for **2b** in TDO). The major differences between the NDO$_{F352V}$ and the TDO$_{F366V}$, regarding the biotransformation of naphthalene, is the complete switch in enantioselectivity of the TDO variant from >98 % *ee* of **2a** to an impressive 90 % *ee* of **2b**. Our findings highlight the relevant role of such mutation to confer access to compound **2b**. Thus, the essential relevance of our work relies on the identification of an engineered TDO variant that is now capable to generate **2b** in substantial quantities in excellent enantiomeric purity. As a result, comparative semi-preparative biotransformations with TDO wild-type and TDO$_{F366V}$, using 20 mM naphthalene, engendered the formation of 20 mM **2a** (100 %) and 6.94 mM **2b** (35 %), respectively. After purification, 287 mg of **2a** (89 %) and 101 mg of **2b** (31 %) were isolated. The structural identity as well as the purity were assessed *via* ^{1}H- and ^{13}C-NMR, chiral HPLC-DAD, and optical rotation measurements.

Comparing the yield of **2a** biosynthesized with TDO wild type, with literature data (Table S1), we achieved with our system not only a quiet high isolated product yield, but also we established an efficient process enabling full conversion. The product titer of 2.87 g L^{-1} for **2a** is in the same range as previously reported (3.33 g L^{-1}), (McIver et al., 2008). Nevertheless, our theoretical yield was far higher, 89 % instead of 7 %. The generation of **2b** with the TDO$_{F366V}$, resulted, to our knowledge, in the first semi-preparative, one-step biosynthesis of this synthon with an excellent enantiomeric excess. We foresee kinetic resolution with

Table 1
Product formation and enantiomeric excess of naphthalene biotransformations with TDO wild-type and variants at position F366.

Catalyst	% *ee*	Product formation [%]
Wild-type	99.6 (+)-(1R,2S)	100.0
F366A	67.6 (-)-(1S,2R)	1.5
F366I	61.4 (-)-(1S,2R)	8.5
F366L	74.1 (-)-(1S,2R)	99.2
F366N	49.8 (-)-(1S,2R)	1.9
F366Q	9.3 (-)-(1S,2R)	36.8
F366S	62.3 (-)-(1S,2R)	3.3
F366T	84.1 (-)-(1S,2R)	17.8
[b]F366V	90.1 (-)-(1S,2R)	80.7
F366Y	[a]nda	[a]nda

[a] nda; no detectable activity.
[b] F366V; Variant TDO$_{F366V}$ outperformed above the rest of the set. This mutant was applied for the semi-preparative biotransformation, enabling the synthesis of **2b** with an excellent enantiomeric excess of 90 % and advantageous product formation.

P. putida 119 or *P. putida* NCIMB 11767A (Allen et al., 1995; Jeffrey et al., 1975), as a viable alternative to eliminate residual traces of **2a** in the TDO$_{F366V}$ driven biosynthesis for **2b**.

We demonstrated the feasibility of the one-step semi-preparative production of **2b** by using the TDO$_{F366V}$ engineered platform. Moreover, our findings provides a solid basis for further studies of TDO behavior towards naphthalene. The effect of F366 variants on naphthalene derivatives, as well as the combination of F366 variants with other TDO hot spot positions is currently under investigation.

CRediT authorship contribution statement

Julian L. Wissner: Conceptualization, Investigation, Methodology, Visualization, Writing - original draft. **Wendy Escobedo-Hinojosa:** Conceptualization, Visualization, Writing - review & editing. **Andreas Vogel:** Methodology, Resources. **Bernhard Hauer:** Funding acquisition, Supervision, Writing - review & editing.

Declaration of Competing Interest

The authors declared no financial or commercial conflict of interest.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.jbiotec.2020.12.007.

References

Allen, C.C.R., Boyd, D.R., Dalton, H., Sharma, N.D., Brannigan, I., Kerley, N.A., Sheldrake, G.N., Taylor, S.C., 1995. Enantioselective bacterial biotransformation routes to cis-diol metabolites of monosubstituted benzenes, naphthalene and benzocycloalkenes of either absolute configuration. J. Chem. Soc. Chem. Commun. 117–118. https://doi.org/10.1039/C39950000117.

Banwell, M.G., Coster, M.J., Edwards, A.J., Karunaratne, O.P., Smith, J.A., Welling, L.L., Willis, A.C., 2003. A total synthesis of the styryllactone (+)-goniodiol from naphthalene. Aust. J. Chem. 56, 585–595. https://doi.org/10.1071/CH02242.

Boyd, D.R., Sharma, D.S., Kerley, N.A., McMoride, R.A., Sheldrake, G.N., Williams, P., Dalton, H., 1996. Dioxygenase-catalysed oxidation of dihydronaphthalenes to yield arene hydrate and cis-dihydro naphthalenediols. J. Chem. Soc. 67–74.

Endoma, M.A., Bui, V.P., Hansen, J., Hudlicky, T., 2002. Medium-scale preparation of useful metabolites of aromatic compounds via whole-cell fermentation with recombinant organisms. Org. Process Res. Dev. 6, 525–532. https://doi.org/10.1021/op020013s.

Gally, C., Nestl, B.M., Hauer, B., 2015. Engineering rieske non-heme iron oxygenases for the asymmetric dihydroxylation of alkenes. Angew. Chem. Int. Ed. 54, 12952–12956. https://doi.org/10.1002/anie.201506527.

Hudlicky, T., Seoane, G., Pettus, T., 1989. Enantioselective synthesis of (-)-zeylena from styrene. J. Org. Chem. 54, 4239–4243. https://doi.org/10.1021/jo00278a052.

Jeffrey, A.M., Yeh, H.J.C., Jerina, D.M., Patel, T.R., Davey, J.F., Gibson, D.T., 1975. Initial reactions in the oxidation of naphthalene by Pseudomonas putida. Biochemistry 14, 575–584. https://doi.org/10.1021/bi00674a018.

Kwit, M., Gawronski, J., Boyd, D.R., Sharma, N.D., Kaik, M., More O'Ferrall, R.A., Kudavalli, J.S., 2008. Toluene dioxygenase-catalyzed synthesis of cis-dihydrodiol metabolites from 2-substituted naphthalene substrates: assignments of absolute configurations and conformations from circular dichroism and optical rotation measurements. Chem. A Eur. J. 14, 11500–11511. https://doi.org/10.1002/chem.200801686.

McIver, A.M., Garikipati, S.V.B.J., Bankole, K.S., Gyamerah, M., Peeples, T.L., 2008. Microbial oxidation of naphthalene to cis-1,2-naphthalene dihydrodiol using naphthalene dioxygenase in biphasic media. Biotechnol. Prog. 24, 593–598. https://doi.org/10.1021/bp070416h.

Orsini, F., Sello, G., Travaini, E., Di Gennaro, P., 2002. A chemoenzymatic synthesis of (2R)-8-substituted-2-aminotetralins. Tetrahedron Asymmetry 13, 253–259. https://doi.org/10.1016/S0957-4166(02)00098-8.

Orsini, F., Sello, G., Bernasconi, S., Fallacara, G., 2004. Chemoenzymatic synthesis of conduritol analogues. Tetrahedron Lett. 45, 9253–9255. https://doi.org/10.1016/j.tetlet.2004.10.072.

Parales, R.E., Resnick, S.M., Yu, C.L., Boyd, D.R., Sharma, N.D., Gibson, D.T., 2000. Regioselectivity and enantioselectivity of naphthalene dioxygenase during arene cis-dihydroxylation: control by phenylalanine 352 in the α subunit. J. Bacteriol. 182, 5495–5504. https://doi.org/10.1128/JB.182.19.5495-5504.2000.

Sello, G., Orsini, F., 2005. Preparation and synthetic use of enantiopure naphthalene dihydrodiols. Mini. Org. Chem. 1, 77–92. https://doi.org/10.2174/1570193043488944.

Vila, M.A., Umpiérrez, D., Veiga, N., Seoane, G., Carrera, I., Rodríguez Giordano, S., 2017. Site-directed mutagenesis studies on the toluene dioxygenase enzymatic system: role of phenylalanine 366, threonine 365 and isoleucine 324 in the chemo-, regio-, and stereoselectivity. Adv. Synth. Catal. 359. https://doi.org/10.1002/adsc.201700444.

Wissner, J.L., Ludwig, J., Escobedo-Hinojosa, W., Hauer, B., 2020. An enhanced toluene dioxygenase platform for the production of cis-1,2-dihydrocatechol in Escherichia coli BW25113 lacking glycerol dehydrogenase activity. J. Biotechnol. https://doi.org/10.1016/j.jbiotec.2020.09.012.

6.4   RESEARCH   ARTICLE   II

Supplementary Information

Supplementary Data

An engineered toluene dioxygenase for a single step biocatalytical production of (-)-(1S,2R)-cis-1,2-dihydro-1,2-naphthalenediol

Julian L. Wissner[a], Wendy Escobedo-Hinojosa[a], Andreas Vogel[b], Bernhard Hauer[a]*

[a]Institute of Technical Biochemistry, University of Stuttgart, Allmandring 31, 70569 Stuttgart, Germany

[b]c-LEcta GmbH, Perlickstr. 5, 04103 Leipzig, Germany

*Corresponding author

E-mail address: bernhard.hauer@itb.uni-stuttgart.de

Supplementary section 1: Material and methods

1.1 Materials

Chemicals and solvents used in this work were obtained at the highest purity degree available from Sigma-Aldrich (St. Louis, US) and Carl Roth (Karlsruhe, DE). Racemic *cis*-1,2-dihydro-1,2-naphthalenediol was purchased at Santa Cruz Biotechnology (Dallas, US). Substances (+)-(1*R*,2*S*)-1,2-dihydro-1,2-naphthalenediol **2a** and (-)-(1*S*,2*R*)-1,2-dihydro-1,2-naphthalenediol **2b** were biosynthesized according to 1.4.

1.2 Bacterial strain, plasmid harboring toluene dioxygenase and mutant library

All biotransformations were performed in a recombinant *E. coli* strain BW25113 harboring toluene dioxygenase (BW25113 pBAD18-TDO), as previously reported (Wissner et al., 2020b). The nine single point variants F366A/I/L/N/Q/S/T/V/Y of the TDO α-subunit were generated with an oligo-based mutagenesis protocol and verified by sequencing by c-LEcta (Leipzig, DE). The previously described pBAD18-TDO plasmid, harboring the TDO wild-type, was used as template.

1.3 Growth conditions, expression, biotransformation and analysis

The TDO wild-type and mutants were grown and expressed as previously described (Wissner et al., 2020a). Analytical biotransformations were performed in 20 mL headspace vials, followed by HPLC-DAD and HPLC-ESI-MS analysis as shown in Figs. S1-S5. Ethanol was employed for the preparation of the substrate-stock (0.5 M naphthalene) and added to a final concentration of 10 mM to the analytical reactions. Chiral HPLC-DAD was performed on an Agilent 1260 Infinity HPLC-DAD system (Santa Clara, US), equipped with normal phase Chiralpak IC column (5 µm, 4.6 x 250 mm, Daicel, Osaka, JP), heated to 30°C. Analysis was performed with a flow rate of 1.4 mL min^{-1} and 10 µL sample injection. The mobile phase was n-hexane/isopropanol (95/5) hold isocratically for 50 min. Peaks were detected at 262 nm (Fig. S7 and S8). The enantiomeric excess (% *ee*) was calculated, using the integral values of (+)-(1*R*,2*S*)- and (-)-(1*S*,2*R*)-*cis*-1,2-dihydro-1,2-naphthalenediol.

1.4 Semi-preparative biosynthesis of (+)-(1*R*,2*S*)- and (-)-(1*S*,2*R*)-*cis*-1,2-dihydro-1,2-naphthalenediol

In order to obtain products **2a** and **2b**, semi-preparative resting cell reactions (0.1 L) were performed. For the production of **2a**, TDO wild-type was used, whereas for the production of **2b** variant TDO$_{F366V}$ was employed. Protein expression and resting cells suspension (0.05 g$_{cww}$ mL^{-1}) were done accordingly to the previously described procedures (Wissner et al., 2020a). Biotransformations were started by the addition of naphthalene (20 mM) suspended in ethanol and incubation at 30°C and 180 rpm for 20 h. The reaction suspension was extracted three times with MTBE (1:1, v/v), and the organic layers were combined. After *in vacuo* evaporation of the organic solvent, the crude reaction mixture was dissolved in 4 mL ACN. The purification step was accomplished by preparative HPLC in an Agilent 1260 Infinity HPLC-DAD system (Santa Clara, US), equipped with a preparative C18-column (Supelco, Discovery C18, 5 μm, 21.2 x 100 mm, Bellefonte, US), coupled to a fraction collector (Agilent 1260 Infinity II, Santa Clara, US). Product purification was performed with a flow rate of

4.0 mL min^{-1} and repeated injections of 100 μL. The mobile phase was water/acetonitrile (30/70) hold isocratically for 19 min. Peaks were detected at 262 nm. Product fractions were collected from minutes 6.0 to 7.5 as depicted in Fig. S6. Finally, the pure product was obtained after the mobile phase was removed *in vacuo*.

1.5 NMR for structural characterization and measurement of the optical rotation

The purity of the biosynthesized products **2a** and **2b** was assessed by ^{1}H and ^{13}C-NMR. Spectra were recorded on a Bruker Avance 500 spectrometer operating at 500.15 and 125.76 MHz, for ^{1}H and ^{13}C, respectively. Spectra were recorded using CD$_3$OD as solvent. The chemical shifts (δ) are reported in parts per million (ppm) relative to the standard tetramethylsilane (TMS, δ = 0). Optical rotations were determined using a Perkin Elmer 241 polarimeter at 589 nm and 22°C.

(+)-(1*R*,2*S*)-*cis*-1,2-dihydro-1,2-naphthalenediol **2a** (287 mg, 88.5%, white solid), biosynthesized with TDO wild-type: ^{1}H-NMR (CD$_3$OD, 500.15 MHz) δ = 7.46 (m, 1 H), 7.24 (m, 2 H), 7.12 (m, 1 H), 6.53 (d, 1 H), 5.99 (dd, 1 H), 4.59 (d, 1 H), 4.30 (m, 1 H) ppm (Fig. S9). ^{13}C-NMR (CD$_3$OD, 125.76 MHz) δ = 137.67, 133.86, 130.22, 129.65, 129.14, 128.90, 128.24, 127.73, 71.91, 69.12 ppm (Fig. 10). $[α]_D^{22}$ = +214 (*c* 0.1, MeOH), Chiral HPLC-DAD: >98% *ee*. All data were in agreement with previously reported references (Jerina et al., 1971; Mukherjee et al., 2010).

(-)-(1*S*,2*R*)-*cis*-1,2-dihydro-1,2-naphthalenediol **2b** (101 mg, 31.1%, white solid), biosynthesized with TDO variant F366V: ^{1}H-NMR (CD$_3$OD, 500.15 MHz) δ = 7.46 (m, 1 H), 7.24 (m, 2 H), 7.12 (m, 1 H), 6.53 (d, 1 H), 6.00 (dd, 1 H), 4.60 (d, 1 H), 4.30 (m, 1 H) ppm (Fig. S11). ^{13}C-NMR (CD$_3$OD, 125.76 MHz) δ = 137.72, 133.89, 130.24, 129.67, 129.19, 128.90, 128.28, 127.73, 71.93, 69.13 ppm (Fig. S12). $[α]_D^{22}$ = -192 (*c* 0.1, MeOH)), Chiral HPLC-DAD: 90% *ee*.

3

Supplementary section 2: Supplementary data

Table S1: Literature compilation reporting the chemical and enzymatic production of (1*R*,2*S*)- and (1*S*,2*R*)-*cis*-1,2-dihydro-1,2-naphthalenediol. Bacterial strain, applied substrate, achieved product yield and the enantiomeric excess are provided. Reactions were performed in semi-preparative scale, unless otherwise marked.

Chemical synthesis / wild-type strain/ recombinant strain (plasmid)	Substrate	Substrate (g / mmol)	Product yield (g / mmol / (%))	Product titer (g L^{-1})	Enantiomeric excess (% *ee*)	Reference
Chemical 3 step synthesis	naphthalene-1,4-dione	1.463 / 9.25	0.720 / 4.44 / (48)	-	0 (*rac*)	(Jeffrey et al., 1974)
Kinetic resolution with *P. putida* 119	(*rac*)-*cis*-1,2-dihydro-1,2-naphthalenediol	-	- / - / -	-	>98 (-)-(1*S*,2*R*)	(Jeffrey et al., 1975)
Kinetic resolution with *P. putida* NCIMB 11767	(*rac*)-*cis*-1,2-dihydro-1,2-naphthalenediol	-	- / - / (40)	-	>98 (-)-(1*S*,2*R*)	(Allen et al., 1995)
[a,b]*P. putida* UV4	1,2-dihydronaphthalene	8.000 / 61.45	2.800 / 17.26/ (28)	0.35	>98 (+)-(1*R*,2*S*)	(Boyd et al., 1996)
[a,b]*P. putida* UV4	1,4-dihydronaphthalene	3.000 / 23.04	[c]3.363 / 20.74 (90)	0.68	>98 (+)-(1*R*,2*S*)	(Boyd et al., 1996)
[d]DH5α (pDTG800)	naphthalene	-	- / - / -	-	40 (+)-(1*R*,2*S*)	(Parales et al., 1998)
[a,b]*E. coli* JM109 (pDTG601)	naphthalene	-	90.000 / 554.94 / -	6.00	>98 (+)-(1*R*,2*S*)	(Endoma et al., 2002)
[e]*E. coli* JM109(DE3) (pDTG141)	naphthalene	2.000 /15.60	[c]0.167 / 1.03 / (7)	3.33	>98 (+)-(1*R*,2*S*)	(McIver et al., 2008)
Chemical 4 step synthesis	2-vinylbenzaldehyde	0.072 / 0.54	0.026 / 0.16 / (29)	-	90 (+)-(1*R*,2*S*)	(Mukherjee et al., 2010)
[b]*E. coli* BW25113 (pBAD18-TDO)	naphthalene	0.256 / 2.00	0.287 / 1.77 / (89)	2.87	>98 (+)-(1*R*,2*S*)	This study
[b]*E. coli* BW25113 (pBAD18-TDO_F366V)	naphthalene	0.256 / 2.00	0.101 / 0.62 / (31)	1.01	90 (-)-(1*S*,2*R*)	This study

a: fermenter; **b**: reaction catalyzed by toluene dioxygenase (TDO); **c**: values report product formation and not isolated yield; **d**: reaction catalyzed by 2-nitrotoluene 2,3-dioxygenase; **e**: reaction catalyzed by naphthalene dioxygenase (NDO); -: not reported
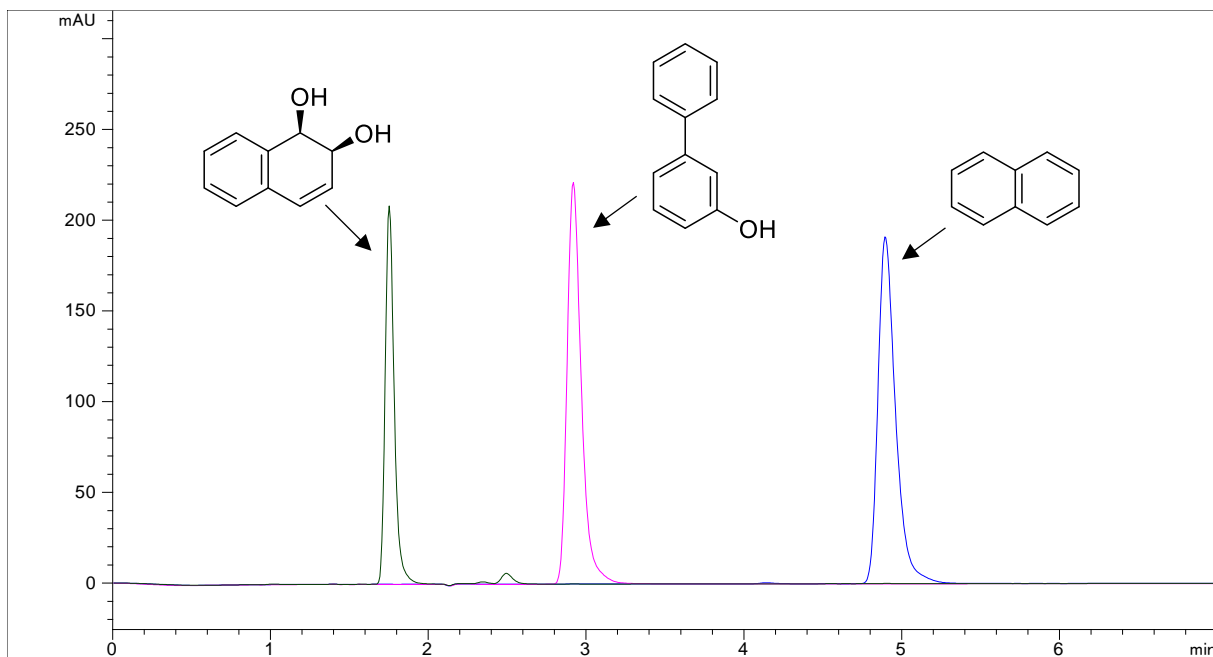
Fig. S1: HPLC-DAD chromatogram (262 nm) of the commercial standard *cis*-1,2-dihydro-1,2-naphthalenediol (green; 1.77 min), the internal standard 3-phenylphenol (pink; 2.92 min) and the substrate naphthalene (blue; 4.90 min).
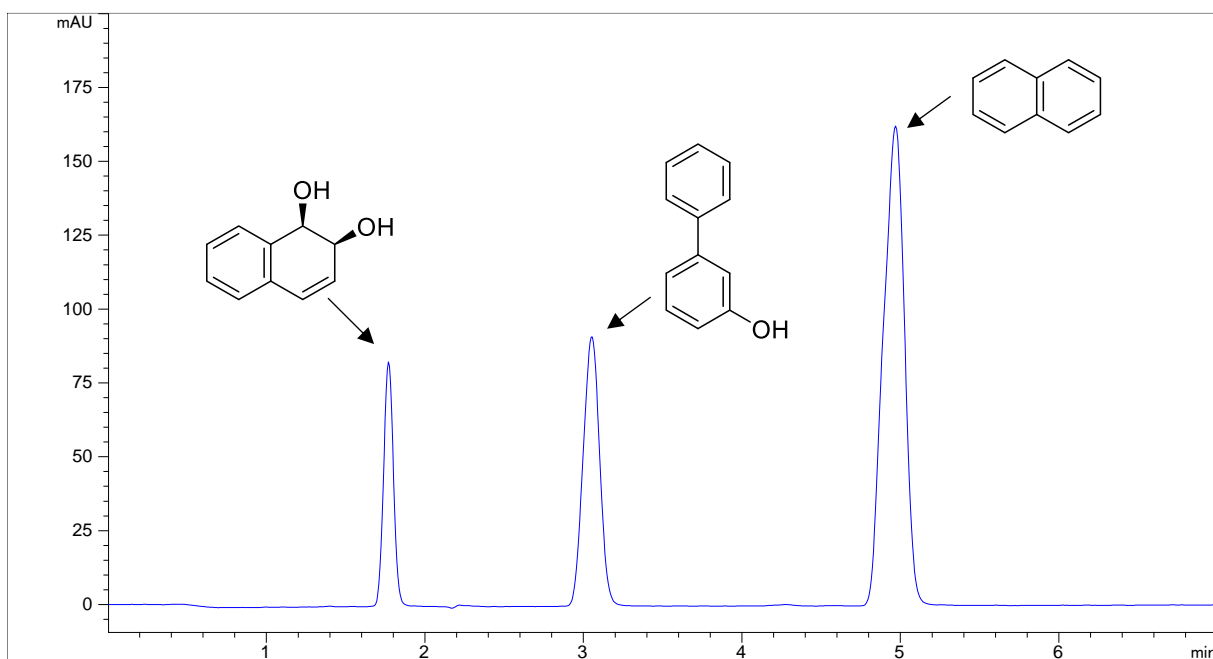

Fig. S2: HPLC-DAD (262 nm) chromatogram of naphthalene biotransformation catalyzed by TDO wild-type after 15 min. *cis*-1,2-Dihydro-1,2-naphthalenediol (1.77 min), internal standard 3-phenylphenol (2.92 min) and naphthalene (4.90 min).
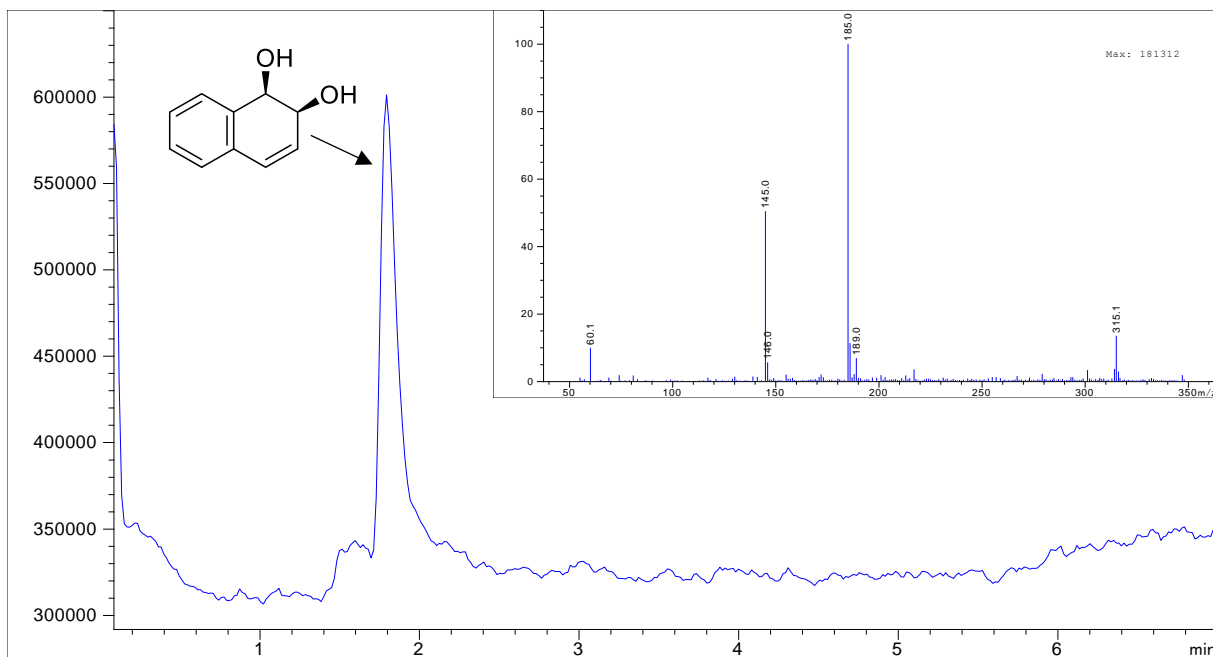
Fig. S3: HPLC-ESI-MS chromatogram of commercial standard *cis*-1,2-dihydro-1,2-naphthalenediol (1.80 min), measured in positive scan. Fragmentation pattern is shown in the right corner (*m/z*: 185 [M+Na], 145 [M+H-H_2O]).
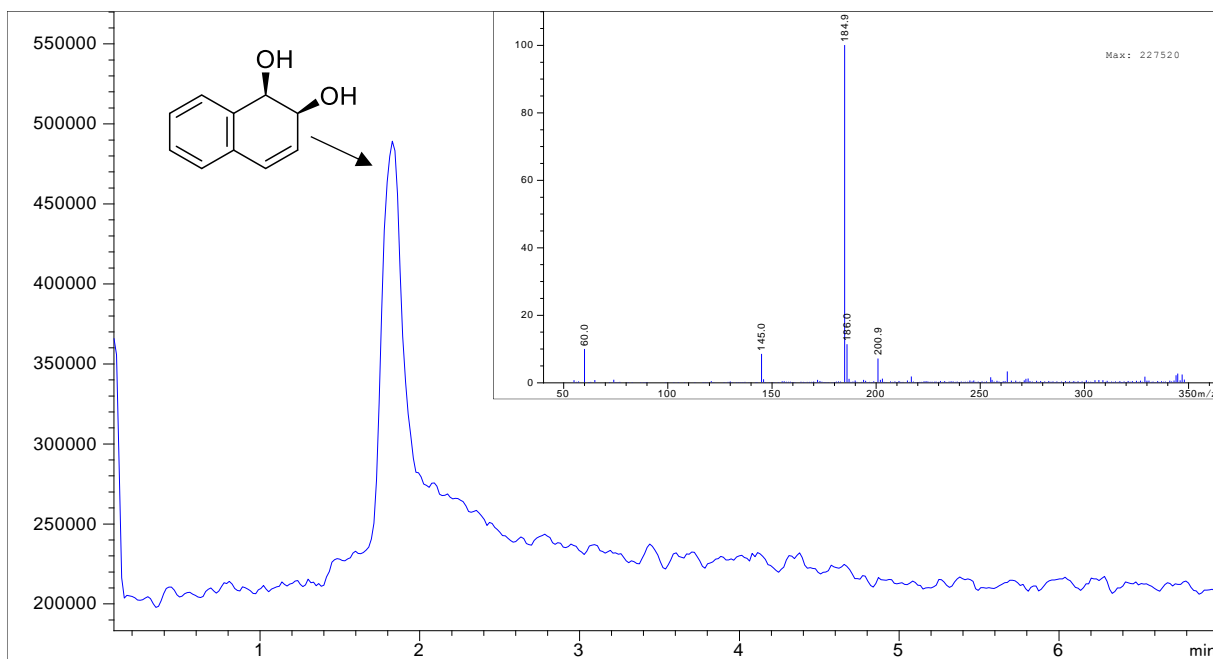
Fig. S4: HPLC-ESI-MS chromatogram of naphthalene biotransformation catalyzed by TDO wild-type, measured in positive scan. *cis*-1,2-Dihydro-1,2-naphthalenediol formation can be seen at 1.80 min. The fragmentation pattern in shown in the right corner (*m/z*: 185 [M+Na], 145 [M+H-H_2O]).

108

6
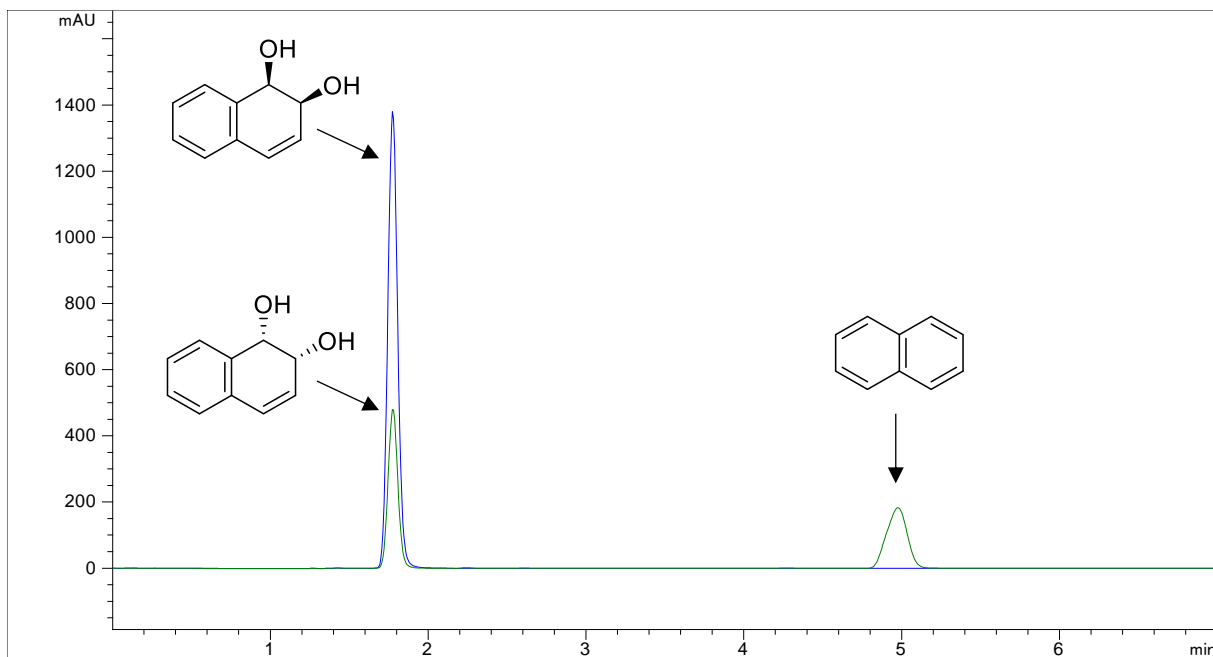
Fig. S5: HPLC-DAD (262 nm) chromatogram of the semi-preparative naphthalene biotransformation (20 mM), catalyzed by; TDO wild-type (blue) and TDO$_{F366V}$ (green). *cis*-1,2-Dihydro-1,2-naphthalenediol enantiomers at 1.77 min and substrate naphthalene at 4.90 min.
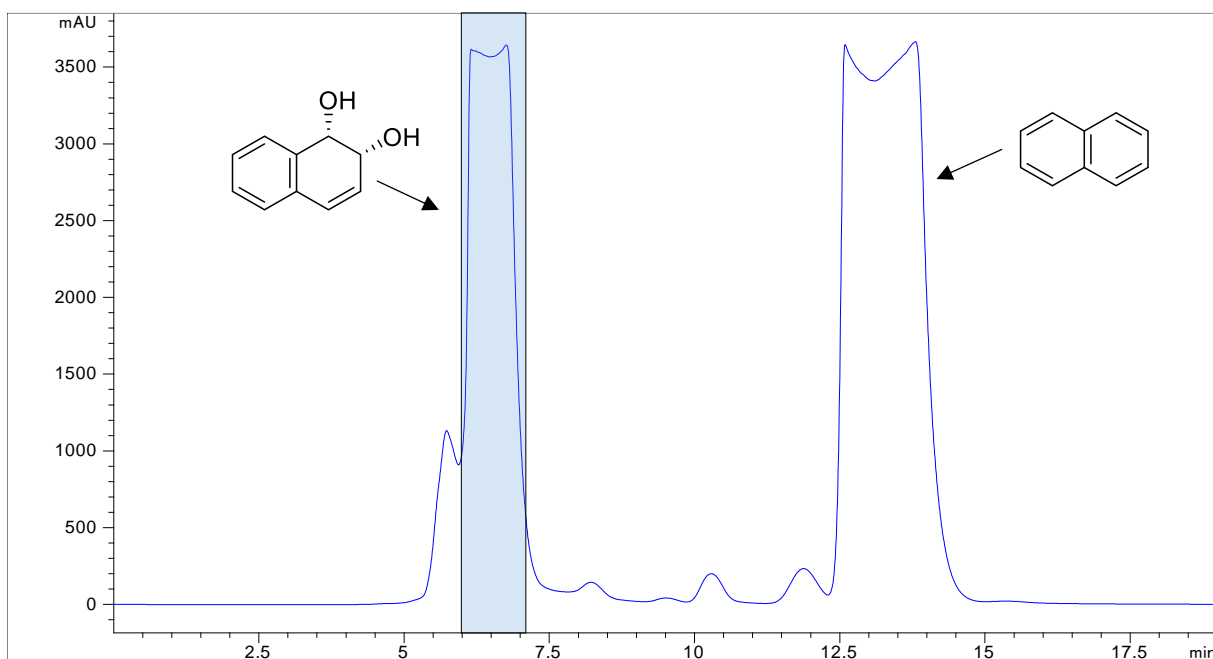
Fig. S6: HPLC-DAD (262 nm) chromatogram of preparative HPLC purification method for naphthalene biotransformation (20 mM) catalyzed by by TDO$_{F366V}$. *cis*-1,2-Dihydro-1,2-naphthalenediol fractions were collected from minute 6.00 to 7.06 (light blue area). Naphthalene substrate was eluted from minute 12.36 to 14.94.
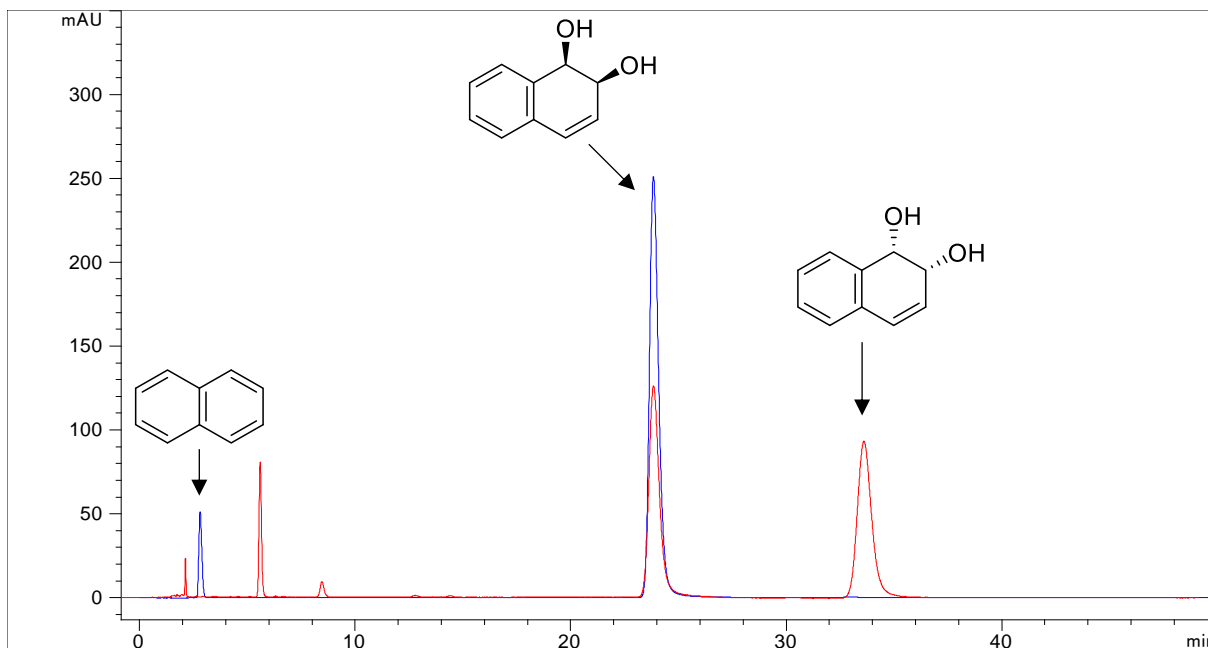
Fig. S7: Chiral HPLC-DAD (262 nm) chromatogram of the commercial available racemic standard (red), and biosynthesized *cis*-1,2-dihydro-1,2-naphthalenediol by TDO wild-type (blue). Recorded retention times; naphthalene 2.78 min, **2a** 23.85 min, and **2b** 33.62 min.
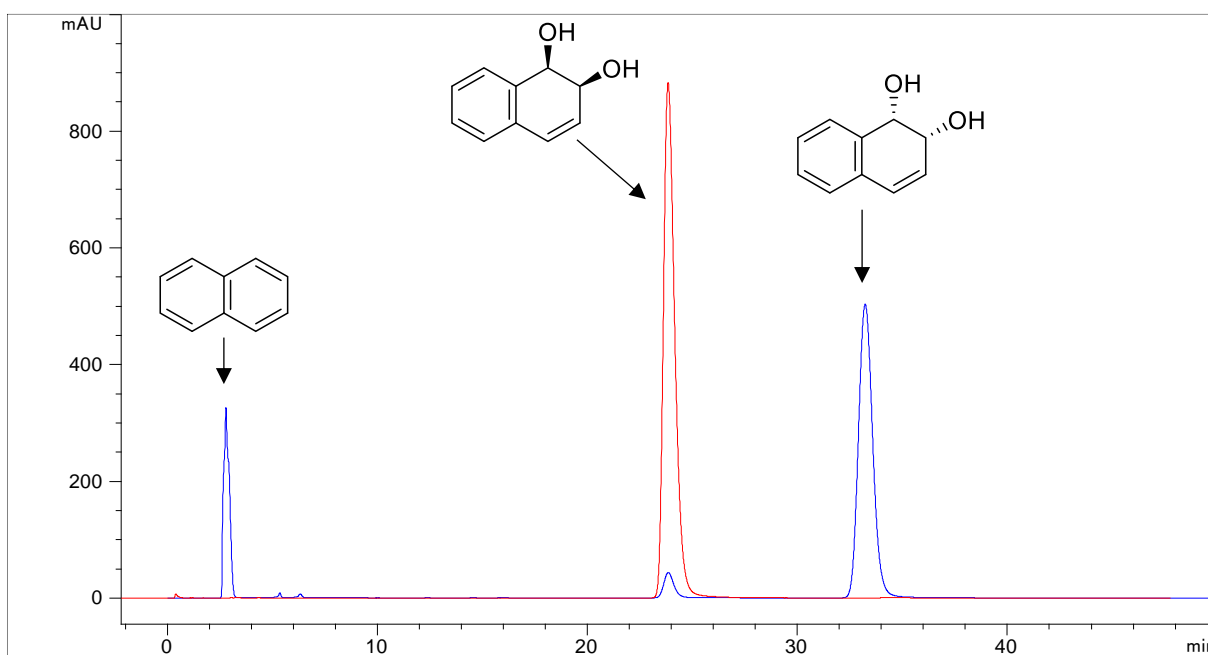
Fig. S8: Chiral HPLC-DAD (262 nm) chromatogram of biosynthesized *cis*-1,2-dihydro-1,2-naphthalenediol employing either TDO wild-type (red), or TDO**F366V** variant (blue). Recorded retention times; naphthalene 2.78 min, **2a** 23.85 min, and **2b** 33.62 min.
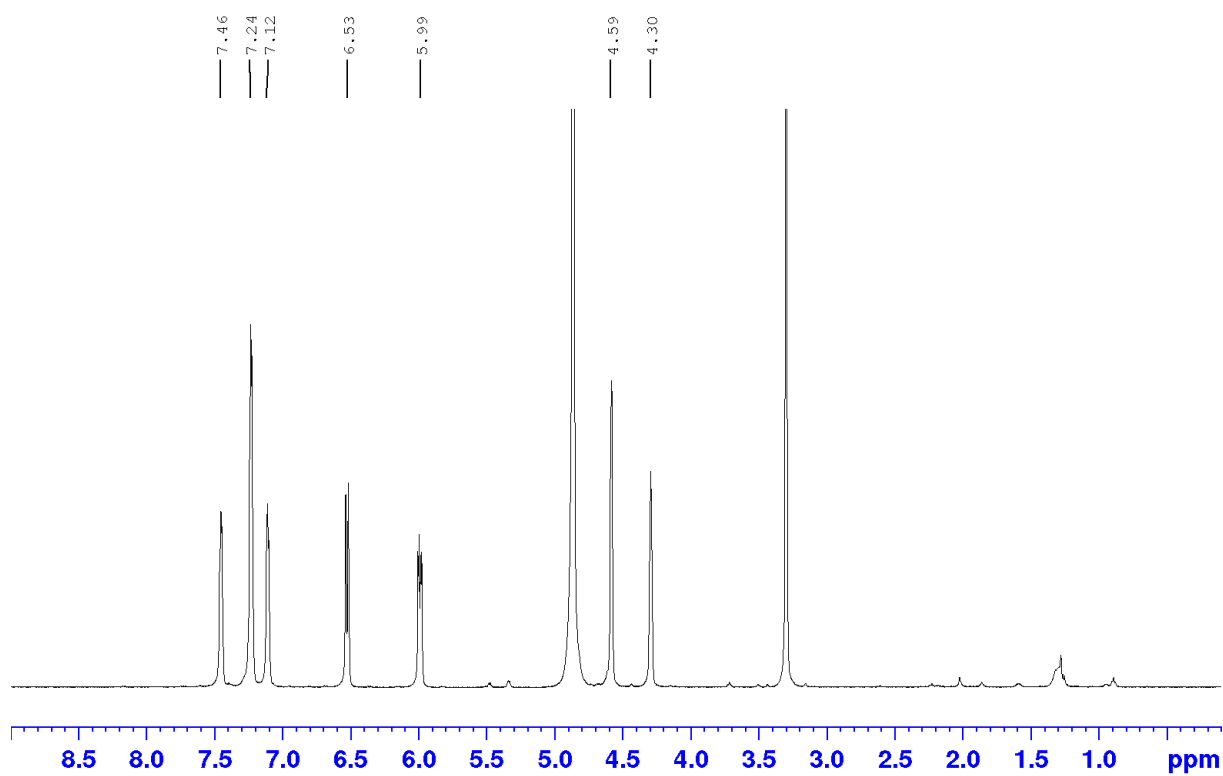
Fig. S9: ¹H-NMR of (+)-(1*R*,2*S*)-*cis*-1,2-dihydro-1,2-naphthalenediol **2a**. H_2O and MeOH can be seen at 4.87 and 3.31 ppm, respectively.
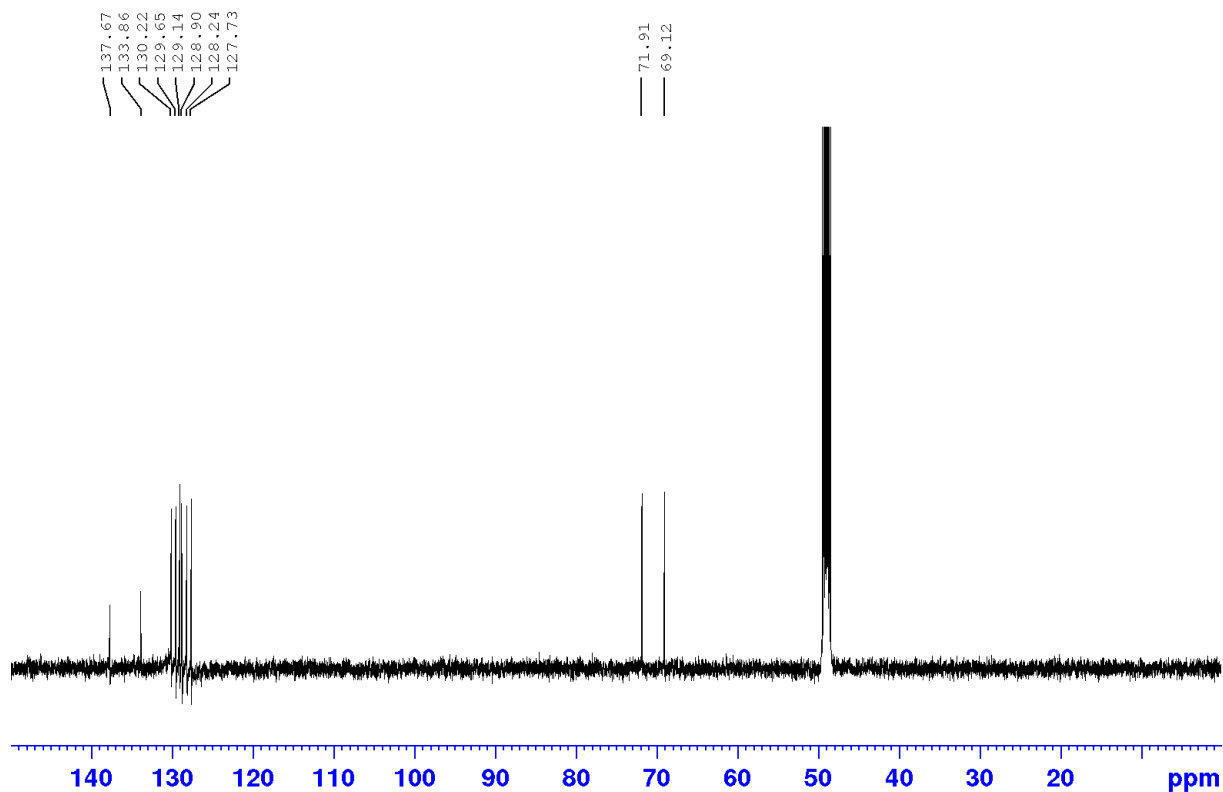
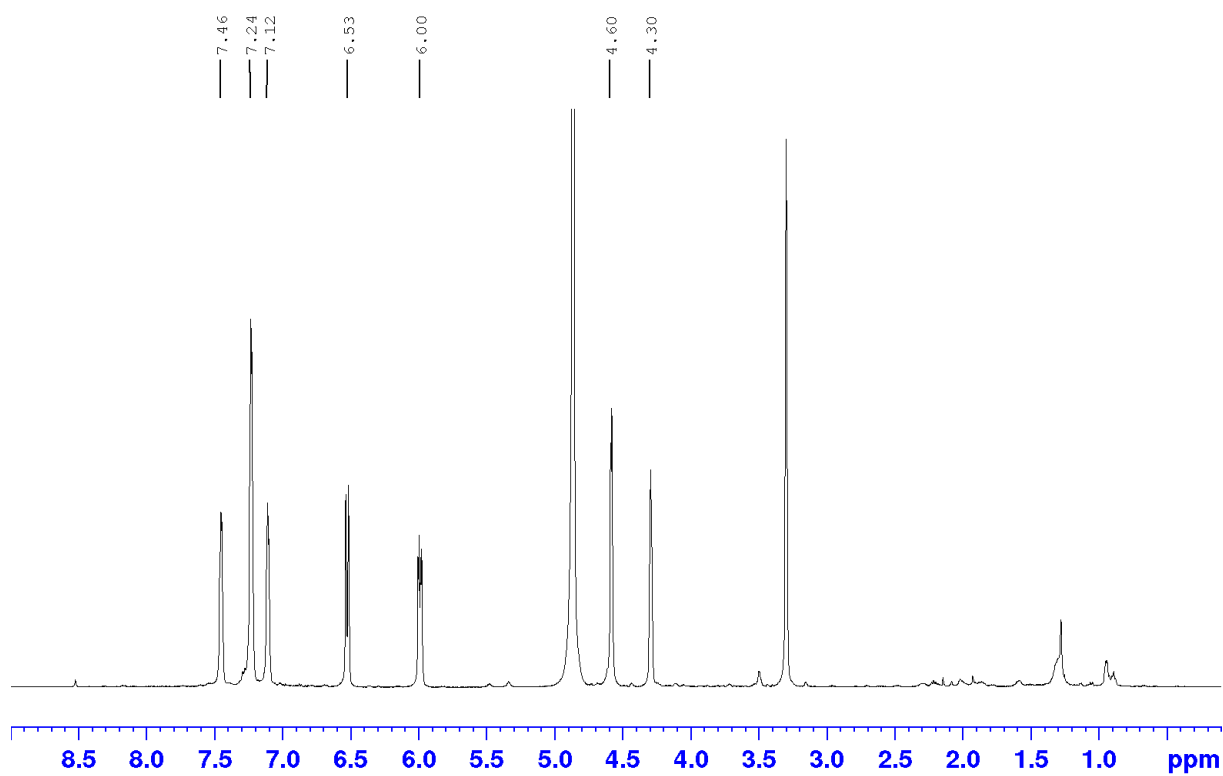Fig. S10: ¹³C-NMR of (+)-(1*R*,2*S*)-*cis*-1,2-dihydro-1,2-naphthalenediol **2a**. MeOH can be seen at 48.06 ppm.

Fig. S11: ^{1}H-NMR of (-)-(1*S*,2*R*)-*cis*-1,2-dihydro-1,2-naphthalenediol **2b**. H$_2$O and MeOH can be seen at 4.87 and 3.31 ppm, respectively.
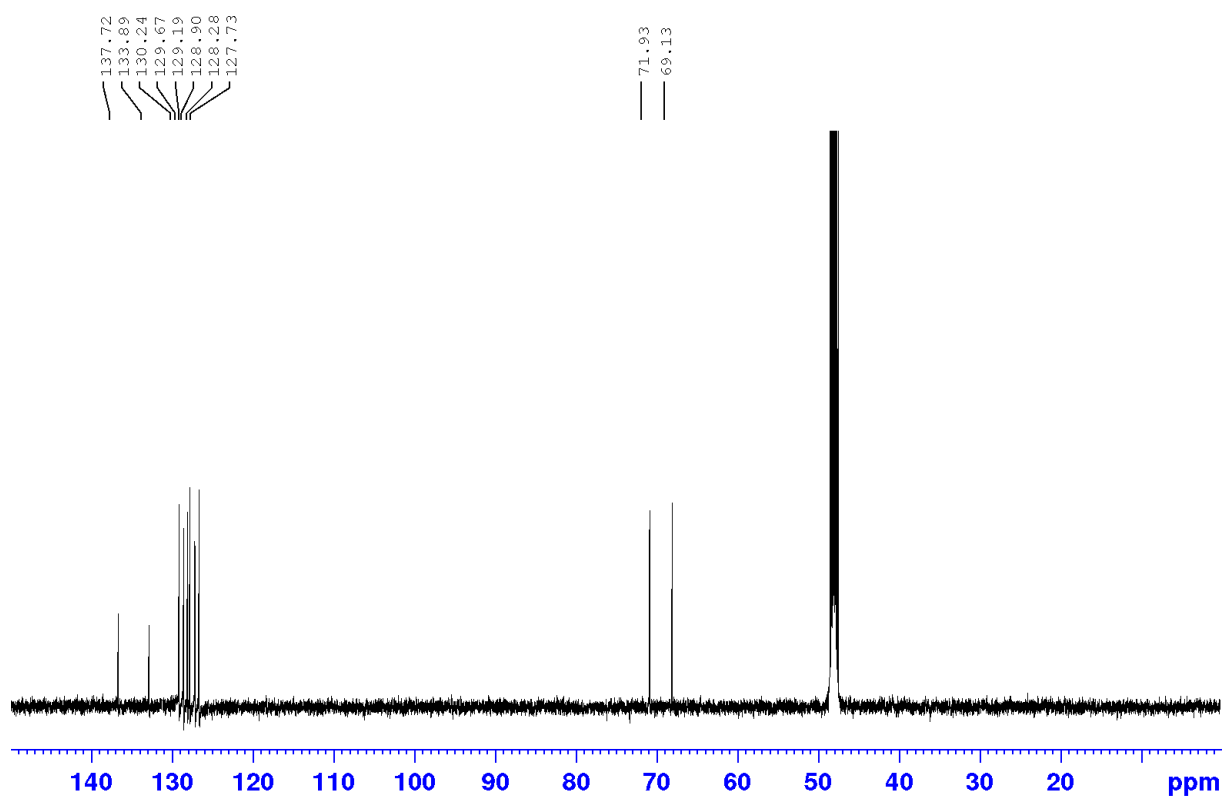

Fig. S12: ^{13}C-NMR of (-)-(1*S*,2*R*)-*cis*-1,2-dihydro-1,2-naphthalenediol **2b**. MeOH can be seen at 48.06 ppm.

Supplementary section 3: Semi-rational design of TDO$_{F366}$ variants

We examined the variants F366A, F366I, F366L, F366T and F366V, due to their influence in enantioselectivity at the corresponding positions in NDO (Parales et al., 2000). In addition we used the 3DM information system of Bio-Prodict (Nijmegen, NL) for alignment of 671 ROs of the three subfamilies; toluene-/benzene, naphthalene and cumene dioxygenase-like, to identify the most frequent amino acids at the position F366 (Fig. S13). The most common amino acid of the RO subset is phenylalanine (71%) followed by glutamine (14%) and valine (4%). Build on the amino acid occurrence in the native dioxygenase subset, in addition the TDO variants F366N, F366Q, F366S and F366Y were examined.
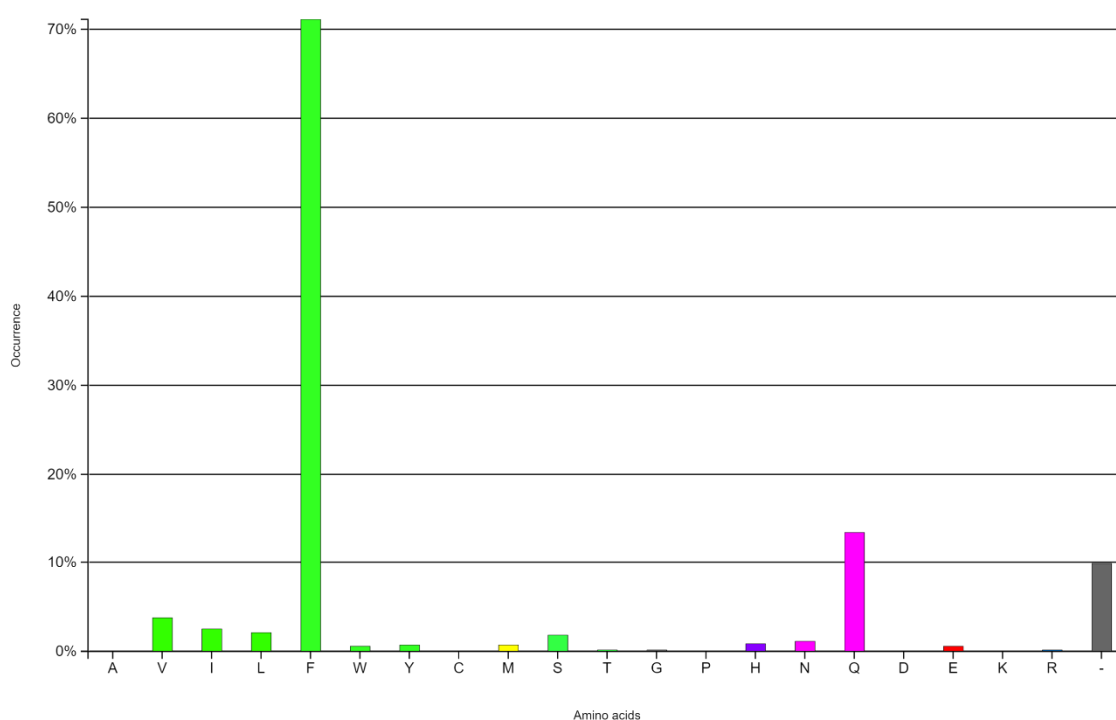

Fig. S13: Amino acid occurrence of 671 Rieske non-heme iron dioxygenases of the three subfamilies; toluene-/benzene, naphthalene and cumene dioxygenase-like, according to the 3DM position 289, corresponding to F366 in toluene dioxygenase from *P. putida* F1. The protein sequence alignment was performed *via* the 3DM information system of Bio-Prodict.

Supplementary section 4: Influence of resting cell amount on naphthalene conversion

We performed experiments using different amounts of resting cells to determine the optimal cell concentration during biotransformations (Fig. S.14). Headspace vials (20 mL) were selected as reaction vessel to provide a 1:20 v/v liquid: air ratio. A single 10 mM addition of naphthalene was supplied to perform the biotransformations. Different cell concentrations (0.010, 0.025, 0.050, 0.100 and 0.200 g_{cww} mL^{-1}), were employed to perform the comparative study.

We observed a linear tendency between the amount of employed cells, from 0.010, up to 0.050 g_{cww} mL^{-1}, and the reached substrate conversion. Employing higher cell concentrations than 0.050 g_{cww} mL^{-1}, lead to a drastic reduction in substrate conversion. Such findings could be explained by the higher consumption of the co-substrate oxygen, by the aerobic respiratory system of whole cells, leaving insufficient oxygen concentrations for full naphthalene conversion. Results highlighted 0.05 g_{cww} mL^{-1} as the optimal cell concentration to achieve full substrate conversion under the tested conditions (100% product formation), and also stressed that lower, as well as higher cell concentrations (0.010, 0.025, 0.100 and 0.200 g_{cww} mL^{-1}) are not really suitable for successful conversions.

In addition, we performed the same reactions in 2 or 5 mL microreaction tubes. Using such smaller reaction containers resulted in a five-fold decrease in substrate conversion, which might be mostly due to the insufficient liquid: air ratio (v/v), (data not shown). These results indicate that whole cell catalyzed biotransformations with the TDO *E.coli* BW25113 platform strongly depend not only on the cell concentration, but also relies in the available liquid: air ratio (v/v). We took into account our findings for the examination of each F366 variant, thus, biotransformations, were performed in 20 mL headspace vials containing 0.05 g_{cww} mL^{-1} and 10 mM of the substrate naphthalene.
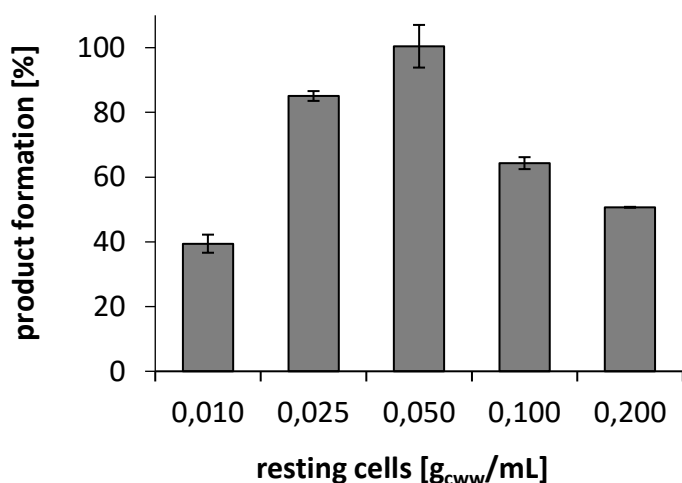

Fig. S14: Achieved *cis*-1,2-dihydro-1,2-naphthalenediol formation, catalyzed by different concentrations of *E. coli* BW25113 resting cells, harboring pBAD18-TDO. Biotransformations were performed with 10 mM naphthalene for 30 min at 30°C, 180 rpm and were stopped by MTBE one-step extraction (1:1 v/v). Data represent the mean value of biological triplicates and their corresponding standard deviations.

References

Allen, C.C.R., Boyd, D.R., Dalton, H., Sharma, N.D., Brannigan, I., Kerley, N.A., Sheldrake, G.N., Taylor, S.C., 1995. Enantioselective bacterial biotransformation routes to cis-diol metabolites of monosubstituted benzenes, naphthalene and benzocycloalkenes of either absolute configuration. J. Chem. Soc. Chem. Commun. 117–118. https://doi.org/10.1039/C39950000117

Boyd, D.R., Sharma, D.S., Kerley, N.A., McMoride, R.A., Sheldrake, G.N., Williams, P., Dalton, H., 1996. Dioxygenase-catalysed oxidation of dihydronaphthalenes to yield arene hydrate and cis-dihydro naphthalenediols. J. Chem. Soc. 67–74.

Endoma, M.A., Bui, V.P., Hansen, J., Hudlicky, T., 2002. Medium-scale preparation of useful metabolites of aromatic compounds via whole-cell fermentation with recombinant organisms. Org. Process Res. Dev. 6, 525–532. https://doi.org/10.1021/op020013s

Jeffrey, A.M., Yeh, H.J.C., Jerina, D.M., 1974. Synthesis of cis-1,2-dihydroxy-1,2-dihydronaphthalene and cis-1,4-dihydroxy-1,4-dihydronaphthalene. J. Org. Chem. 39, 1405–1407. https://doi.org/10.1021/jo00926a018

Jeffrey, A.M., Yeh, H.J.C., Jerina, D.M., Patel, T.R., Davey, J.F., Gibson, D.T., 1975. Initial Reactions in the Oxidation of Naphthalene by Pseudomonas Putida. Biochemistry 14, 575–584. https://doi.org/10.1021/bi00674a018

Jerina, D.M., Daly, J.W., Jeffrey, A.M., Gibson, D.T., 1971. Cis-1,2-dihydroxy-1,2-dihydronaphthalene: A bacterial metabolite from naphthalene. Arch. Biochem. Biophys. 142, 394–396. https://doi.org/10.1016/0003-9861(71)90298-0

McIver, A.M., Garikipati, S.V.B.J., Bankole, K.S., Gyamerah, M., Peeples, T.L., 2008. Microbial oxidation of naphthalene to cis-1,2-naphthalene dihydrodiol using naphthalene dioxygenase in biphasic media. Biotechnol. Prog. 24, 593–598. https://doi.org/10.1021/bp070416h

Mukherjee, P., Roy, S.J.S., Sarkar, T.K., 2010. A diversity-oriented synthesis of bicyclic cis -dihydroarenediols, cis -4-hydroxyscytalones, and bicyclic conduritol analogues. Org. Lett. 12, 2472–2475. https://doi.org/10.1021/ol100557f

Parales, R.E., Resnick, S.M., Yu, C.L., Boyd, D.R., Sharma, N.D., Gibson, D.T., 2000. Regioselectivity and enantioselectivity of naphthalene dioxygenase during arene cis-dihydroxylation: Control by Phenylalanine 352 in the α subunit. J. Bacteriol. 182, 5495–5504. https://doi.org/10.1128/JB.182.19.5495-5504.2000

Parales, J. V., Parales, R.E., Resnick, S.M., Gibson, D.T., 1998. Enzyme specificity of 2-nitrotoluene 2,3-dioxygenase from Pseudomonas sp. strain JS42 is determined by the C-terminal region of the α subunit of the oxygenase component. J. Bacteriol. 180, 1194–1199. https://doi.org/10.1128/jb.180.5.1194-1199.1998

Wissner, J.L., Escobedo-Hinojosa, W., Heinemann, P.M., Hunold, A., Hauer, B., 2020a. Methods for the detection and analysis of dioxygenase catalyzed dihydroxylation in mutant derived libraries, in: Enzyme Engineering and Evolution, Including Directed Evolution. Elsevier Inc., pp. 1–31. https://doi.org/10.1016/bs.mie.2020.04.022

Wissner, J.L., Ludwig, J., Escobedo-Hinojosa, W., Hauer, B., 2020b. An enhanced toluene dioxygenase platform for the production of cis-1,2-dihydrocatechol in Escherichia coli BW25113 lacking glycerol dehydrogenase activity. J. Biotechnol. https://doi.org/10.1016/j.jbiotec.2020.09.012

6.5 RESEARCH   ARTICLE   III

Manuscript

Advanced
Synthesis &
Catalysis

Check for updates

Semi-Rational Engineering of Toluene Dioxygenase from *Pseudomonas putida* F1 towards Oxyfunctionalization of Bicyclic Aromatics

Julian L. Wissner,[a] Jona T. Schelle,[a] Wendy Escobedo-Hinojosa,[a] Andreas Vogel,[b] and Bernhard Hauer[a],*

[a] Institute of Technical Biochemistry, University of Stuttgart, Allmandring 31, 70569 Stuttgart, Germany
     Tel: +49 711 685–63193;
     E-mail: bernhard.hauer@itb.uni-stuttgart.de
[b] c-LEcta GmbH, Perlickstr. 5, 04103 Leipzig, Germany

Supporting information for this article is available on the WWW under https://doi.org/10.1002/adsc.202100296

Abstract: Toluene dioxygenase (TDO) from *Pseudomonas putida* F1 was engineered towards the oxyfunctionalization of bicyclic substrates. Single and double mutant libraries addressing 27 different positions, located at the active site and entrance channel were generated. In total, 176 different variants were tested employing the substrates naphthalene, 1,2,3,4-tetrahydroquinoline, and 2-phenylpyridine. Introduced mutations in positions M220, A223 and F366, exhibited major influences in terms of product formation, chemo-, regio- and enantioselectivity. By semi-rational evolution, we lighted up the TDO capability to convert bulkier substrates than its natural substrate, at unprecedented reported conversions. Thus, the most active TDO variants were applied to biocatalytic oxyfunctionalizations of 1,2,3,4-tetrahydroquinoline, and 2-phenylpyridine, enabling the production of substantial amounts of (+)-(*R*)-1,2,3,4-tetrahydroquinoline-4-ol (71% isolated yield, 94% *ee*) and (+)-(1*S*,2*R*)-3-(pyridin-2-yl)cyclohexa-3,5-diene-1,2-diol (60% isolated yield, 98% *ee*), respectively. Here, we provide a set of novel TDO-based biocatalysts useful for the preparation of oxyfunctionalized bicyclic scaffolds, which are valuable to perform downstream synthetic processes.
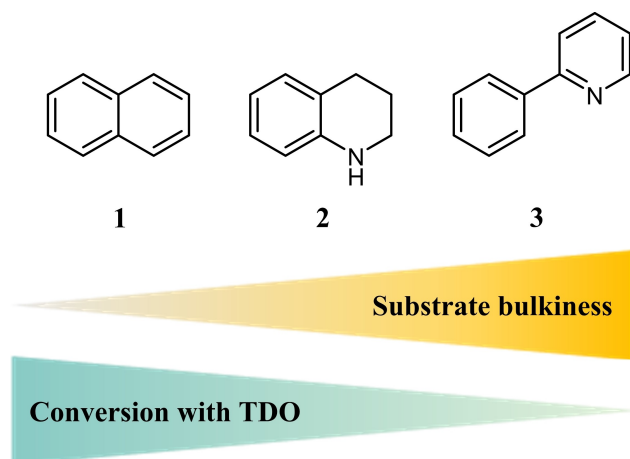
Keywords: Biocatalysis; Biotransformations; Hydroxylation; Mutagenesis; Rieske non-heme iron dioxygenases

Introduction

The selective oxyfunctionalization of aromatic and heteroaromatic compounds is of great interest for the generation of chiral synthons for pharmaceuticals and fine chemicals.[1–4] Mono- and dihydroxylated aromatics can be obtained by the use of multicomponent Rieske-non heme iron dioxygenases (ROs), which consist of an oxygenase, a reductase and optionally a ferredoxin. ROs are capable to incorporate molecular oxygen into aromatics, creating valuable and chiral *cis*-dihydrodiendiols. Furthermore, this remarkable class of enzymes is capable to perform a variety of different reactions, for instance monohydroxylations, sulfoxidations, desaturations and C-H amination.[5–8] Toluene dioxygenase (TDO) from *Pseudomonas putida* F1, is one of the best studied ROs in the last decades. Its striking promiscuity embraces over 100 non-natural substrates.[9] In terms of catalytic efficiency, TDO is capable of affording astonishingly high titers of monocyclic aromatics, such as bromobenzene, at up to $35 \, \mathrm{g \, L^{-1}}$.[10] Nevertheless, in comparison to naphthalene dioxygenase (NDO), cumene dioxygenase (CDO), and biphenyl dioxygenase (BPDO), from of *Pseudomonas putida* 9816-4, *Pseudomonas fluorescens* IP01, and *S. yanoikuyae* B8/36, respectively, TDO yields lower product formation for bicyclic aromatic substrates. Such reduced performance of TDO in front of bicyclic

aromatics might correlate with the smaller active pocket size, when compared with the ones found in NDO and BPDO.[11] For instance, for the flat bicyclic substrate naphthalene, TDO generates 10 g L^{-1} of the dihydroxylation product, while for the more sterically demanding biphenyl it yields only 3.0 g L^{-1}.[1] Besides, the presence of one nitrogen atom in one of the aromatic rings, results in a drastic decrease in TDO product formation, such as in the case of 2-phenyl-pyridine (only 1%).[11] This pattern suggests that TDO displays a substrate-size dependent performance (Scheme 1). In order to overcome limitations regarding product formation and selectivity of TDO, previous studies employed directed evolution techniques for generating TDO variants.[12,13] For instance, Zhang and colleagues evolved TDO using random mutagenesis towards the conversion of indene to *cis*-indandiol, a potential key intermediate in the chemical synthesis of the anti-viral Crixivan®.[14] A drawback of directed evolution is the massive screening effort that has to be performed, usually comprising thousands of variants. An attractive and feasible alternative is the rational-design of a reduced set of TDO variants via site-directed mutagenesis. In this case, by employing either bioinformatics tools, such as, *in silico* studies or previous knowledge of experimental data, candidate key positions can be identified and utilized for mutagenesis.[15–17] For instance, Vila and co-workers employed a smartly designed set of TDO active site variants at positions I324, T365 and F366, all three exhibiting major influences in regio- and enantioselectivity.[18] To investigate hot-spot positions

influencing the conversion of bicyclic aromatics, we designed a set of semi-rational TDO variants, which were generated via site-directed mutagenesis by c-LEcta (Leipzig, DE). Our approach resulted in the identification of engineered TDO variants displaying remarkable conversion activities towards all tested bicyclic aromatics. Findings reported herein foster the biocatalytic synthesis of valuable oxyfunctionalized aromatic bicyclic scaffolds.

Results and Discussion

To our best knowledge, a detailed study on the influence of the TDO active site residues, driving product formation, chemo-, regio- and stereoselectivity was so far, not performed. In order to foster the TDO catalyzed conversion of bicyclic aromatic substrates, we employed site-directed mutagenesis aiming to engineering; 1) the active site, 2) the putative substrate entrance channel, and 3) literature known key positions of the TDO system (supplementary section 3).

1) For the generation of active site variants, we used the 3DM database of Bio-Prodict (Nijmegen, NL), to align the sequences of 671 ROs of the toluene/ benzene, naphthalene, and cumene dioxygenase sub-families. All active site positions[15,19] with a conservation below 90% were chosen for mutagenesis (Q215, F216, M220, A223, G264, Y266, L272, I276, V309, H311, L321, I324, F366 and F372). The highly conserved position H222, H228 and D376, coordinating the catalytic iron, as well as D219 were therefore not considered. The 14 active site positions were mutated into amino acids, which exhibited according to the sequence alignment, an occurrence of over 2% (Table S2).

2) The substrate channel was modelled using the crystal structure of the TDO α-subunit (PDB ID; 3EN1) and the CAVER plugin (version 3.0.1) for PyMOL (supplementary section 6). To investigate the influence of the substrate channel size, all amino acids around 5 Å of the putative substrate channel (G224, T225, L229, L245, P247 and P248) were replaced by an alanine (Table S3).

3) At last, seven positions, either previously described in literature or suggested by the 3DM database as hot-spot position (F114, T115, A212, A234, G323, V340, T365), were chosen for mutagenesis.[14,18,20,21]

Thus, a total of 121 single variants at 27 different amino acid positions were generated (Table S1) and investigated towards the bicyclic substrates; naphthalene **1**, 1,2,3,4-tetrahydroquinoline **2**, and 2-phenylpyridine **3** (Scheme 1). Lastly, since the screening of the mutant library with substrates **1**–**3** highlighted variants at positions M220 and A223 as beneficial for an increased product formation, we selected such positions to generate a second mutant

Scheme 1. Substrates employed in this study for the characterization of TDO catalyzed hydroxylation patterns. Naphthalene **1**, 1,2,3,4-tetrahydroquinoline **2**, and 2-phenylpyridine **3**. Substrate bulkiness (yellow gradient), increases from left **1** to right **3**. Substrate conversion displayed by TDO wild type (green gradient), shows an inverse correlation with respect to the substrate bulkiness from left **1** (high conversion) to right **3** (low conversion).

119

library consisting out of 55 double variants focused on positions M220 and A223 (Table S4). Consequently, we describe in the current study 176 semi-rational designed single and double variants (Table S1 and S4). In addition to the characterization of these variants, we aimed for the semi-preparative production of the identified products, employing the outperforming engineered TDO variants (Scheme 2).

Screening of the mutant library was performed according to experimental section 1.3. Variants fulfilling one or more of the following criteria were validated according to experimental section 1.4 as follows; 1) Variants displaying over 200% total product formation in comparison with the TDO wild type, 2) Variants showing a switch in product distribution in favor to the side product, and in addition, at least 50% total product formation, when compared to TDO wild type, 3) Variants generating selectively one product, 4) Double variants have to be at least as active or as selective as the parent single variant. In addition, the enantiomeric excess for **1a**, **2a**, and **3a** was determined via chiral HPLC-DAD for the validated variants (experimental section 1.9).

The screening of the single mutant library with substrate **1**, highlighted the active site positions M220 and A223 boosting the formation of **1a** (Figure S25). Validation of the single point variants M220A, A223I and A223V showed a 2.6-, 2.3-, and 5.0-fold increase in product formation, in comparison with TDO wild

type, respectively (Table 1). The enantiomeric excess for **1a** ($>98\%$), remained the same for all three variants. Double variants at the positions M220 and A223 were not able to surpass the product formation of their corresponding single point parent variants. Thus, for **1a** the highest product formation of 6.27 ± 0.08 mM (62.7%) was achieved with the variant A223V, employing 10 mM substrate, within only 0.5 h, clearly illustrating the feasibility of engineering TDO towards bicyclic aromatics. As previously reported, active site variant F366V exhibited a switch from $>98\%$ *ee* **1a** to -90, while decreasing, in comparison to the TDO wild type, 20.8-fold the product formation.[17] Such dramatic reduction in naphthalene conversion with TDO_{F366V} compared to the fast occurring TDO wild type driven reaction, can be explained by the short established reaction time (only 0.5 h) and the restrictive working conditions, in terms of air to liquid ratio, to perform the biotransformations. Double variants TDO_{F216A_F366V} and TDO_{M220A_F366V} showed nor a further improvement in enantioselectivity (-70% *ee* and 13% *ee*, respectively), neither an increased product formation, compared to the single point variant TDO_{F366V} (data not shown).

TDO wild type showed an initial acceptance for substrate **2**, yielding products **2a** and **2b** after 20 h of biotransformation employing 10 mM of **2**. The mono-hydroxylated compound **2a** was the main product (2.74 ± 0.18 mM, 77% *ee*), while the dehydrogenated compound **2b** was secondary (0.22 ± 0.02 mM). As expected, the observed conversion for **2** (only 30%) was lower in comparison to **1** (100%), due to the larger size of substrate **2**. It is worth to mention that product **2a** is of great interest, considering its relevance as pharmaceutical synthon.[22–24] Thus, fostering **2a** production became an enticing task for this work. By testing the single point TDO library with substrate **2** we identify the active site positions A223 and L321, and position F114, located at the putative ferredoxin binding site,[20] as positive hits favoring the total product formation (Figure S26). Validation of the variants F114H and A223V indicated a 1.9- and 1.7-fold increase in total product formation, respectively,
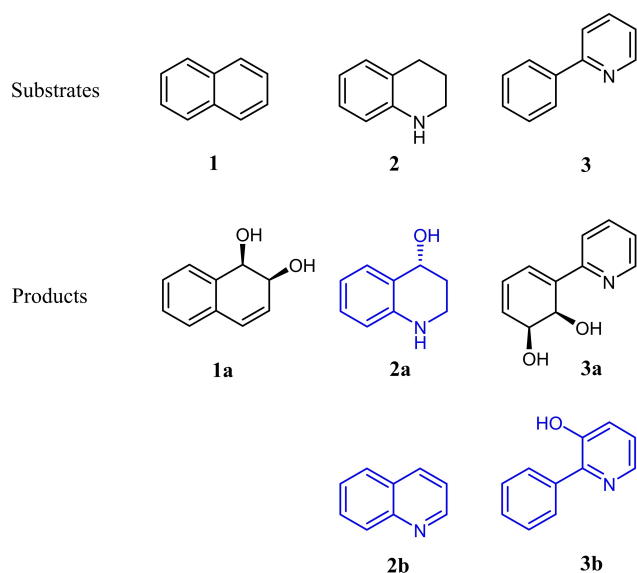
Scheme 2. TDO wild type catalyzed conversion of the substrates naphthalene **1**, 1,2,3,4-tetrahydroquinoline **2**, and 2-phenylpyridine **3**. Generated products; (+)-(1*R*,2*S*)-1,2-dihydro-1,2-naphthalenediol **1a**, (+)-(*R*)-1,2,3,4-tetrahydroquinoline-4-ol **2a**, quinoline **2b**, (+)-(1*S*,2*R*)-3-(pyridin-2-yl)cyclohexa-3,5-diene-1,2-diol **3a**, and 2-phenylpyridin-3-ol **3b**. Novel products identified in this work, **2a**, **2b**, and **3b**, are depicted in blue.

Table 1. Validation of variants with increased product formation and/or switched product distribution for naphthalene **1** (10 mM).

Variant	PF of **1a** [mM][a]	*ee* of **1a** [%][b]
WT	1.25 ± 0.04	>98
M220A	3.28 ± 0.24	>98
A223I	2.85 ± 0.69	>98
A223V	6.27 ± 0.08	>98
F366V	0.06 ± 0.03	-90

[a] PF stands for product formation.
[b] *ee* stands for enantiomeric excess.

Adv. Synth. Catal. **2021**, *363*, 1–11   **Wiley Online Library**   3

These are not the final page numbers! ↗↗

© 2021 The Authors. Advanced Synthesis & Catalysis published by Wiley-VCH GmbH

and maintained the same product distribution (**2a**:**2b**; 92:8), as for the TDO wild type driven reaction (Table 2). Active site variant A223T was not only capable of converting the total amount of substrate **2** (10 mM), but also favored the synthesis of the interesting product **2a** (**2b** concentration remained < 0.04 mM). Variant L321A also increased the total product formation 1.9-fold, however, the exhibited product distribution was altered (**2a**:**2b**; 72:28). In contrast, active site variants M220A, L321S and F366V showed a switch in chemoselectivity, favoring now the generation of the secondary product **2b**. The observed distributions **2a**:**2b** were as follows; 28:72, 17:83 and, 26:74, for variants M220A, L321S, and F366V, respectively. Along with the chemoselectivity switch, a decreased total product formation was observed in comparison with TDO wild type, resulting in a reduction of 3.3-, 2.4-, and 3.1-fold, for mutants M220A, L321S, and F366V, respectively. Double variant TDO$_{F114H\_A223T}$ surpassed the beneficial chemoselectivity of the single variant TDO$_{A223T}$, since full conversion was achieved, and product formation was exclusively directed toward **2a** formation. In addition, the combinatorial double variant F114H_A223T exhibited higher enantiomeric excess (83% *ee*), compared to the wild type and the corresponding single parent variants F114H and A223T (for both cases, 77% *ee*). Thus, the double variant not only outperformed TDO wild type, but also all tested variants, enabling the selective production of 9.93 ± 0.24 mM (99.3%) of **2a** with 83% *ee*. Interestingly, all single and double variants containing amino acid exchanges at position M220, exhibited, compared to TDO wild type, a decreased product formation along with an altered chemoselectivity. We managed to foster the enantiomeric excess in favor to **2a** by altering the pH value of the reaction buffer (supplementary section 5). Thus, at

higher pH-value the enantiomeric excess for **2a** was increased. Such effect might be explained by the racemization of **2a** at acidic conditions.[25] As a result, biotransformations performed in 0.1 M TRIS-buffer pH 8.5 exhibited an *ee* of 97 and 95%, for TDO wild type and for TDO$_{F114H\_A223T}$, respectively. Moreover, the **2a**:**2b** product ratio remained the same as observed for the biotransformations performed in 0.1 M potassium phosphate buffer pH 7.4 (Table S8). An increased total product formation (4.13 ± 0.46 mM) was obtained in the case of TDO wild type driven reactions employing TRIS-buffer pH 8.5. However, the opposite effect was observed for TDO$_{F114H\_A223T}$, resulting in a product reduction down to 8.58 ± 0.18 mM, instead of the 9.93 ± 0.24 mM generated product observed in potassium phosphate buffer pH 7.4. Among all tested variants, TDO$_{F366V}$ showed the highest total product formation favoring **2b** (2.79 ± 0.11 mM, **2a**:**2b**; 27:73), in biotransformations performed in TRIS-buffer pH 8.5. Besides, altering TRIS-buffer pH to 9.0, led to even higher *ee* values for **2a**, but at the expense of total product formation (data not shown). For a conclusive product identification and structural characterization, semi-preparative biotransformations in 0.1 M TRIS-buffer pH 8.5 were performed, employing the selected variants TDO$_{F114H\_A223T}$ and TDO$_{F366V}$, for the generation of **2a** and for **2b**, respectively. Semi-preparative biotransformation of **2** with TDO$_{F114H\_A223T}$ yielded 106 mg (71%, 94% *ee*) of **2a** as sole product, while TDO$_{F366V}$ yielded 20 mg (16%) of **2b** and minor amounts of **2a** (< 5 mg; not isolated).

In agreement with literature, conversion of 2-phenylpyridine **3** with TDO wild type yielded **3a** as main product, in slight amounts (0.50 mM, 5%)[11] Such TDO conversion behavior confirmed the trend illustrated in Scheme 1, showing an inverse correlation in terms of substrate size and achieved conversion. Naturally, not only the bulkiness, but also the electronic properties of the substrates can play a role, especially for electron-poor heteroaromatics. Strikingly, TDO conversion of **3** generated, in addition to compound **3a**, the secondary product **3b** (0.11 mM, 1%), a chemical entity not previously reported in the literature. The discovery of **3b** turned out to be of great interest, since the enzymatic hydroxylation of a heteroaromatic moiety in a molecule, connected to an aromatic system, was so far not observed.[11] Electron-poor heteroaromatics, for instance pyridine, are much more resistant to RO catalyzed hydroxylation, than aromatics.[26] Usually, RO catalyzed monohydroxylation of alkyl groups attached to heteroaromatics and *cis*-dihydroxylation of fused aromatics, occur both more readily than the hydroxylation of heteroaromatics. The formation of the monohydroxylated product **3b** involves most likely an unstable *cis*-dihydrodiendiol which undergoes spontaneous dehydration as described for the TDO catalyzed hydroxylation of 4-

Table 2. Validation of variants with increased product formation and/or switched product distribution for 1,2,3,4-tetrahydroquinoline **2** (10 mM) at pH 7.4 (0.1 M potassium phosphate buffer).

Variant	PF of **2a** [mM][a]	PF of **2b** [mM][a]	*ee* of **2a** [%][b]
WT	2.74 ± 0.18	0.22 ± 0.02	77
F114H	5.07 ± 0.40	0.44 ± 0.04	77
M220A	0.25 ± 0.01	0.64 ± 0.02	60
A223T	9.91 ± 0.21	0.04 ± 0.01	77
A223V	4.61 ± 0.17	0.49 ± 0.02	78
L321A	3.96 ± 0.20	1.55 ± 0.07	70
L321S	0.21 ± 0.04	1.03 ± 0.09	69
F366V	0.25 ± 0.01	0.73 ± 0.04	83
F114H_A223T	9.93 ± 0.24	0.01 ± 0.01	83

[a] PF stands for product formation.
[b] *ee* stands for enantiomeric excess.

Adv. Synth. Catal. **2021**, *363*, 1–11   **Wiley Online Library**   4

These are not the final page numbers! ↗↗

© 2021 The Authors. Advanced Synthesis & Catalysis published by Wiley-VCH GmbH

picoline to 3-hydroxy-4-picoline.[12] Such novel finding of **3b**, encouraged us to boost its production. Therefore, we investigate whether one of our generated variants was capable to favor **3b** product formation to achieve its substantial synthesis.

The screening of the single point mutant library with substrate **3** highlighted the active site positions M220, A223 and V309 as beneficial for increasing the total product formation (Figure S27). In comparison to TDO wild type, validation of variants M220A, M220C, A223I, A223V and V309G showed a 12.4-, 8.5-, 6.5-, 6.4-, and 9.4-fold increase in the total product formation, respectively (Table 3). However, among all the positive hits, only variant M220A maintained the product ratio (**3a**:**3b**; 82:18), observed in TDO wild type, whereas variants A223I and A223V favored the production of **3a** (**3a**:**3b**; 91:9 and 86:14, respectively). In contrast, variant M220C exhibited an increase in product formation for **3b**, by altering the product distribution (**3a**:**3b**; 64:36). Thought variants F216A and F366V, compared to the wild type, exhibited a decreased total product formation (4.4- and 8.7-fold, respectively), both were capable to astoundingly switch the product ratio in favor to **3b** (**3a**:**3b**; 12:88 and 0:100, respectively). Despite we focused our attention on the **3b** entity, we were also interested in following the library performance toward **3a** synthesis. In this way, we found a strong improvement in **3a**

product formation by employing the combinatorial double variants TDO$_{M220A\_A223V}$ and TDO$_{M220A\_V309G}$. Both combinations led to an impressive 14.0- and 15.1-fold increase in total product formation, compared to the wild type, both favoring the production of the *cis*-dihydrodiendiol **3a** (**3a**:**3b**; 92:8 and 93:7, respectively). Product formation increase might be correlated with the enlargement of the active pocket, for variant TDO$_{M220A\_V309G}$ (Figure S31). Regarding the enantiomeric excess for product **3a**, no alterations were detected since all tested variants showed an *ee* > 98%. For practical purposes and a viable production of substantial amounts of both **3a** and **3b**, we selected TDO$_{M220A\_V309G}$ to perform a semi-preparative biotransformation. This approach enabled the formation of 8.51 ± 0.31 mM (85.1%) **3a** and 0.64 ± 0.03 mM (6.4%) **3b**, resulting in a conversion of 91.5%, which is impressive considering that the starting performance of TDO wild type was as low as 6.1%. In terms of isolated product yield, this translates in 114 mg (60%; *ee* > 98%) and 6 mg (4%) for **3a** and **3b**, respectively.

We identified semi-rational engineered TDO variants, showing enhanced product formation and selectivity towards bulkier bicyclic aromatic compounds in comparison to the already well accepted substrate naphthalene (Scheme 3).[17] TDO$_{F114H\_A223T}$ and TDO$_{M220A\_V309G}$ enabled **2** and **3** conversion up to 100 and 91.5% respectively, employing 10 mM of substrate, in comparison with the starting point 29.6 and 6.1%, for TDO wild type. These two variants enabled the production and isolation of the highly valuable compounds **2a** and **3a** with remarkable isolated yields (71% and 60%, respectively), advantageous product selectivity (**2a**:**2b**; 100:0, **3a**:**3b**; 93:7), and outstanding enantiomeric excess (94% and > 98% *ee*, respectively). The seven TDO positions, influencing either chemo-, regio- or enantioselectivity, as well as the product formation are summarized in Table 4. All of these positions, except F114, are located in the active site of the TDO (Figure S29).[15,19] Position F114 (corresponding to F99 in NDO from *Pseudomonas sp.* NCIB 9816-4) is located at the ferredoxin binding site, near the Rieske center.[20] Alteration of positions M220, A223 and F366 had a positive impact in terms of chemo-, regio-, enantioselectivity or product formation, for all three tested substrates.

Table 3. Validation of variants with increased product formation and/or switched product distribution for 2-phenylpyridine **3** (10 mM).

Variant	PF of **3a** [mM][a]	PF of **3b** [mM][a]	*ee* of **3a** [%][b]
WT	0.50 ± 0.08	0.11 ± 0.02	> 98
F216A	0.02 ± 0.01	0.13 ± 0.02	n.d.
M220A	6.26 ± 0.13	1.32 ± 0.06	> 98
M220C	3.32 ± 0.21	1.89 ± 0.16	> 98
A223I	3.62 ± 0.28	0.37 ± 0.03	> 98
A223V	3.36 ± 0.46	0.54 ± 0.08	> 98
V309G	5.50 ± 0.40	0.24 ± 0.01	> 98
F366V	0.00 ± 0.00	0.07 ± 0.01	n.d.
M220A_A223V	7.87 ± 0.12	0.64 ± 0.15	> 98
M220A_V309G	8.51 ± 0.31	0.64 ± 0.03	> 98

[a] PF stands for product formation.
[b] *ee* stands for enantiomeric excess.

Table 4. TDO positions with major influence on chemo-, regio- and enantioselectivity and positions exhibiting, in comparison to the wild type, an increased product formation for substrates; naphthalene **1**, 1,2,3,4-tetrahydroquinoline **2**, and 2-phenylpyridine **3**.

	TDO positions with major influence on:			
Substrate	Chemoselectivity	Regioselectivity	Enantioselectivity	Product formation
1	–	–	F366	M220, A223
2	M220, A223, L321, F366	–	–	F114, A223, L321
3	–	F216, M220, V309, F366	–	M220, A223, V309
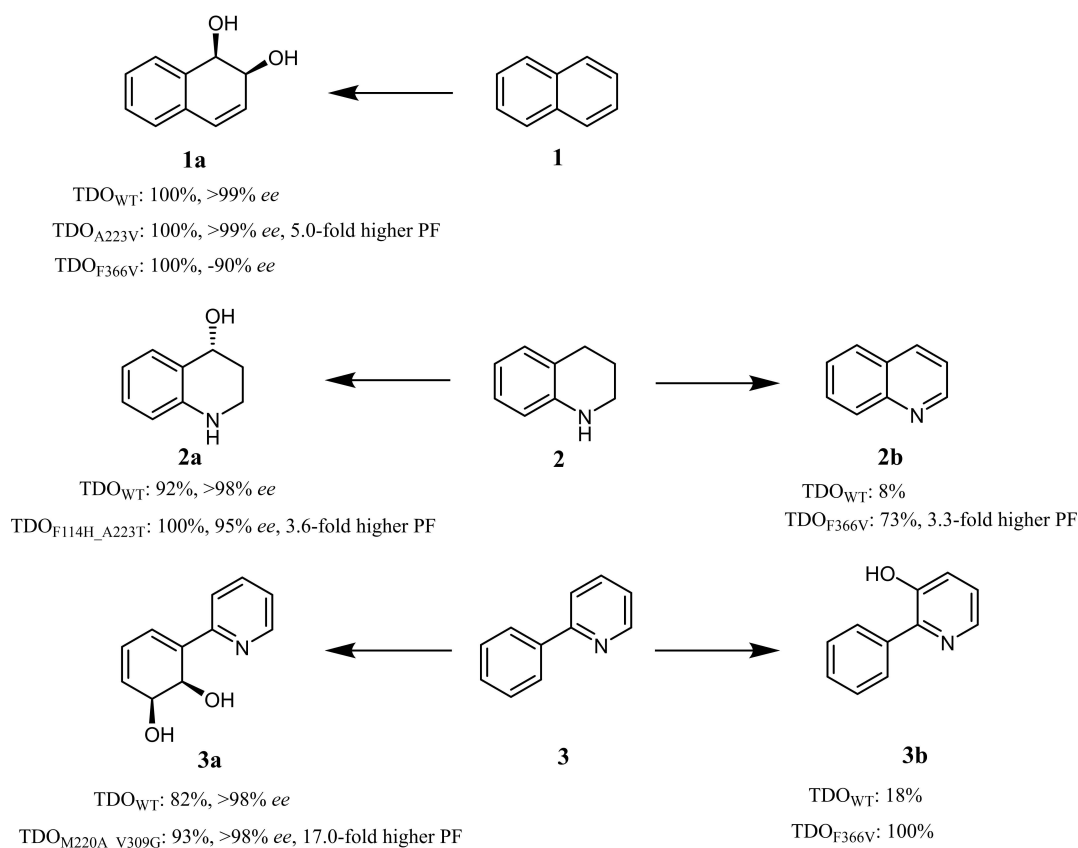
Scheme 3. TDO wild type (TDO$_{WT}$) and identified TDO variants showing enhanced product formation and selectivity, towards the bicyclic substrates **1**, **2** and **3**. Product selectivity (in %) and enantiomeric excess (where suitable, in %), for the wild type and the best variants, is depicted below the generated products for each tested substrate. Higher increase in product formation (PF) for outperforming variants is given in −fold terms, in comparison to the wild type.

Variants at active site position TDO M220, exhibited quite diverse properties regarding the converted substrate. For instance, for substrates **1** and **3**, M220 variants increased the total product formation. Moreover, in the case of **3**, introduced changes also exerted an influence in terms of regioselectivity. All tested M220 variants exhibited an alteration in chemoselectivity for substrate **2**, which mostly favored **2 b** production, highlighting the impact of this position. Influence of TDO M220 position on product formation and selectivity has been described in literature, reporting that variant TDO$_{M220A}$ enabled activity towards 1,2,4,5-tetrachlorobenzene, which was not observed with TDO wild type.[27] Relevance of position M220 was also reported for homolog dioxygenase systems such as NDO and CDO, from *Pseudomonas sp.* NCIB 9816-4 and *Pseudomonas fluorescens* IP01, respectively. For NDO, the exchange of alanine to valine at the equivalent position A206 showed substantial influence on product formation, chemo-, regio- and stereoselectivity for substrates ethylbenzene, α-methylstyrene, and allylbenzene.[28] Whereas for CDO, an exchange of methionine to alanine at the corre-

sponding position M232 enabled, for the first time, the conversion of (+)-α-pinene and also exhibited for a variety of substrates, a positive influence in product formation, regio- and enantioselectivity.[29] In this work, for TDO position M220, the exchange to the small amino acids alanine and cysteine led to an increase in product formation for substrates **1** and **3**. The favorable effect of the introduced changes suggest that it could be correlated with the enlargement of the active pocket.

TDO variants addressing position A223 led to an increase in product formation for all three tested substrates. In the particular case of **2**, variants at position A223 exhibited an influence in chemoselectivity. For **3**, the exchange into the small amino acids valine and threonine, as well as to the aliphatic amino acid isoleucine fostered product formation, resulting in a conversion rise up to 6.5-fold. This is the first time that the mentioned catalytic advantages of TDO variants at position A223 are reported. Besides, despite the beneficial role of this position, it has been not described for any of the related TDO homologs and their equivalent positions, such as NDO from *Pseudo-*

monas sp. NCIB 9816-4 (position V209), BPDO from *Pseudomonas pseudoalcaligenes* KF707 (position A234), or CDO from *Pseudomonas fluorescens* IP01 (position A235). The only data available addressing equivalent TDO position A223 was reported for aniline dioxygenase (ADO) from *Acinetobacter sp.* strain YAA. In this case, the exchange into an alanine of the corresponding position V205 in ADO, induced two different effects; on one hand a novel activity, enabling the conversion of 2-isopropylaniline, and on the other hand, a severe conversion decrease for substrates aniline and 2,4-dimethylaniline.[30] Thus, our approach allowed the identification of TDO position A223 as a novel hot spot location strongly influencing product formation.

TDO position F366 showed perhaps, the most versatile influence. Particularly, variant F366V is of great interest, since in reference to the wild type this exchange prompted three different effects; 1) a complete switch in enantioselectivity for **1 a** (from > 98% to −90% *ee*), 2) a switch in chemoselectivity towards **2 b** (from 92:8 **2 a**:**2 b** to 27:73), and 3) a complete switch in regioselectivity towards **3 b** (from 82:18 **3 a**:**3 b** to 0:100). However, compared to the wild type, all variants at position F366 exhibited a reduced product formation. The influence on enantioselectivity of position TDO F366 was described for naphthalene,[17] bromobenzene, styrene and indene.[18] In the case of the later, in addition to the enantioselectivity influence an alteration in regioselectivity was also observed. Variants at the corresponding position in NDO from *Pseudomonas sp.* NCIB 9816-4 (position F352), exhibited alterations in terms of enantioselectivity for naphthalene, biphenyl, and phenanthrene and also in regioselectivity for the former and the later.[31,32] BPDO variants from *Pseudomonas pseudoalcaligenes* KF707 at the corresponding position (F377) showed altered chemo- and regioselectivity for various chlorinated biphenyls.[33] Taking into account all the reports, it seems that the conserved phenylalanine at position 366 in TDO is not only critical to direct product enantioselectivity in this dioxygenase system, but also in the homologue enzymes NDO and BPDO.

Favorably, for **3**, the bulkier substrate, product formation and product distribution were both enhanced by generating engineered variants which were designed based on a combinatorial approach of the best performing single variants. This approach fostered beyond the observed activity for the single point variants which were already improved in comparison with TDO wild type, for instance, in the case of TDO$_{M220A\_V309G}$ driven conversion of **3**. Interestingly, despite of the enlargement generated by exchanging the amino acids located at the putative substrate entrance channel into alanine (Table S3, Figure S30), no beneficial effects were observed.

Conclusion

In the present study we identified a total of seven positions in TDO, critically influencing the oxyfunctionalization of substrates **1**, **2**, and **3**. Furthermore, the hot spot positions M220, A223, and F366 exhibited for all of three tested substrates substantial changes in product formation and in chemo-, regio- and enantioselectivity. Single and double TDO variants at these positions enabled a conversion over 91% out of 10 mM of **2** and **3**. Conversion achievements for **2** and **3** are outstanding, considering that both substrates were initially poorly converted by the wild type enzyme. In addition, semi-preparative biotransformations employing outperforming variants enabled the TDO catalyzed synthesis of products **2 a**, **2 b**, **3 a**, and **3 b**, for the first time, in substantial amounts. In conclusion, the semi-rational engineering of TDO indeed endowed a viable oxyfunctionalization towards bulky bicyclic aromatics, in terms of product formation, chemo-, regio- and stereoselectivity.

Experimental Section

Materials. Chemicals and solvents used in this work were obtained at the highest purity degree available from Sigma-Aldrich (St. Louis, US) and Carl Roth (Karlsruhe, DE). 2-Phenylpyridin-3-ol **3 b** was purchased at Fluorochem (Hadfield, UK). Substances **2 a**, **2 b**, **3 a**, and **3 b** were biosynthesized according to the details given in 1.5.

Bacterial strain, plasmid harboring toluene dioxygenase and mutant library. All biotransformations were performed in a recombinant *E. coli* strain BW25113 harboring the multicomponent toluene dioxygenase (BW25113 pBAD18-TDO), as previously reported.[34] A number of 176 single- and double variants of the TDO α-subunit were generated by c-LEcta (Leipzig, DE). The pBAD18-TDO plasmid, harboring the TDO wild type, was used as template and variants were generated following an oligo-based mutagenesis protocol and verified by Sanger sequencing.

TDO mutant library screening. TDO mutant libraries consist of pBAD18-TDO variants harbored in recombinant *E. coli* BW25113 cells, which were received in the form of glycerol stocks from c-LEcta (Leipzig, DE). Libraries were constantly stored at −80 °C. Cells were reactivated in 96 deep well plates by inoculating 2 μL of the glycerol stocks in 0.5 mL of LB medium supplemented with ampicillin (100 mg L^{-1}). Plates were covered with Breathe-EASIER sealing membranes (Sigma, St. Louis, US) and incubated overnight at 37 °C and 800 rpm in a microplate shaker (VWR International, Radnor, US). Each overnight culture (25 μL) was transferred into a new 96 deep well plate containing, per well, 1.0 mL of TB medium supplemented with ampicillin (100 mg L^{-1}) and were incubated for 2 h at 37 °C and 800 rpm. Subsequently, protein production was induced by addition of L-arabinose to a final concentration of 10 mM, and plates were incubated at 25 °C at 800 rpm for 20 h. Afterwards, cells were harvested by centrifugation at 4000 × *g* at 4 °C for 20 min and the supernatant was discarded.

Adv. Synth. Catal. **2021**, *363*, 1–11   **Wiley Online Library**   7

These are not the final page numbers! ↗↗

© 2021 The Authors. Advanced Synthesis & Catalysis published by Wiley-VCH GmbH

Cell pellets were resuspended in 0.1 M potassium phosphate buffer (0.49 mL, pH 7.4) supplemented with 20 mM glucose monohydrate and transferred into 96 deep well plates furnished with glass Hirschmann vials (Hirschmann Laborgeräte GmbH & Co. KG, Eberstadt, DE), enabling reusability of 96 deep well plates without cross contamination. Reactions were started by addition of 10 µL of a 0.5 M substrate solution prepared in ethanol, to a final concentration of 10 mM. Biotransformations were incubated at 30 °C and 800 rpm, in the case of substrate naphthalene **1**, only for 0.5 h and in the case of substrates 1,2,3,4-tetrahydroquinoline **2** and 2-phenylpyridine **3** for 20 h. Reactions were stopped by addition of 0.5 mL of the extraction solvent MTBE. After vigorous mixing and centrifugation (4000 × g at room temperature for 5 min) the organic layer was removed and analyzed via HPLC-DAD, to quantify the achieved product formation and selectivity. In addition, HPLC-ESI-MS analysis was performed to confirm the masses of the generated products. The whole set of screening results is shown in supplementary section 4, Figures S25–S27.

Validation of mutant library hits. Variants fulfilling one or more of the following criteria were reevaluated with the corresponding substrate: 1) Variants displaying over 200% total product formation in comparison with TDO wild type, 2) Variants showing a switch in product distribution in favor to the side product, and in addition, at least 50% total product formation, when compared to TDO wild type, 3) Variants generating selectively one product, 4) Double variants have to be at least as active or as selective as the parent single variant and fulfill at least one of the criteria previously mentioned.

Biotransformations were performed according to 1.3, with the only difference, that no glass Hirschmann inlets were used. This increased the total working volume from 1.0 mL to 2.0 mL, resulting in an increased product formation for all tested substrates. In addition, the enantiomeric excess of the generated products was determined via chiral HPLC-DAD. Validation results of variants fulfilling the established criteria are shown in Table 1, 2 and 3.

Semi-preparative biotransformations. Semi-preparative biotransformations were performed according to 1.4 with the only exception that for substrate 1,2,3,4-tetrahydroquinoline **2**, the TDO$_{F114H\_A223T}$ and TDO$_{F366V}$ catalyzed biotransformations were performed in TRIS-buffer (0.1 M, pH 8.5) supplemented with 20 mM glucose monohydrate instead of potassium phosphate buffer (see supplementary section 5). For semi-preparative biotransformation of **3**, TDO$_{M220A\_V309G}$ was employed. A number of 200 simultaneous driven biotransformations were performed in 96 deep well plates, and combined after reaction completion, thus representing a total of 100 mL reaction volume. The combined reaction mixtures were extracted with MTBE either five times (1:1) for **2**, or nine times (1:2) for **3**. The combined organic layers were evaporated *in vacuo* and redissolved in 6 mL of an acetonitrile-water mixture. The purification step was performed by preparative HPLC in an Agilent 1260 Infinity HPLC-DAD system (Santa Clara, US), equipped with a preparative C18-column (Supelco, Discovery C18, 5 µm, 21.2 × 100 mm, Bellefonte, US) and a fraction collector (Agilent 1260 Infinity II, Santa Clara, US). Product purification was performed with a flow rate of 4.0 mL min⁻¹ and repeated injections of 70 µL of the crude reaction mixture. The

mobile phase for the purification of generated products from substrates **2** and **3** was water/acetonitrile (55/45), which was hold isocratically for 22 min. Peaks were detected at 210 nm for the reaction analysis of **2** and at 310 nm for **3** (Supplementary section 1, Figs. S6 and S13). The collected fractions of each product were evaporated *in vacuo* and analyzed via ¹H- and ¹³C-NMR (supplementary section 2, Figs. S14–S24).

Chemical synthesis of (-)-(S)-1,2,3,4-Tetrahydroquinoline-4-ol. The chemical synthesis of (−)-(*S*)-1,2,3,4-tetrahydroquino-line-4-ol was done according to Kolcsar and collegues.[35] The raw product was extracted five times with MTBE, evaporated *in vacuo* and dissolved in 4 mL ACN. Preparative HPLC purification was performed as described in 1.5. A total of 28.1 mg (10% yield) (−)-(*S*)-1,2,3,4-tetrahydroquinoline-4-ol was obtained. The enantiomeric excess for raw and purified product was determined via chiral HPLC-DAD to 81% *ee* and 35% *ee*, respectively.

HPLC-DAD for quantification. For substrate **1** biotransformations HPLC-DAD and ESI-MS methods were performed according to Wissner and colleagues.[17] For the analysis of the biotransformation products of substrates **2** and **3**, an Agilent 1260 Infinity II system (Santa Clara, US), equipped with a C18-column (Supelco C18 Discovery, 5 µm, 4.0 × 150 mm, Bellefonte, US) and a diode array detector (Agilent 1260 Infinity II DAD HS, Santa Clara, US) was operated isothermally at 30 °C and 50 °C for **2** and **3**, respectively. Measurements were run at a flow rate of 1.0 mL min⁻¹. For **2**, the mobile phase was water/methanol 55/45 (v/v) hold isocratically for 10 min and 1 µL sample was injected. For **3**, the mobile phase was water/methanol with a linear gradient of; t = 0 min, 55/45 (v/v); t = 4.0 min, 50/50 (v/v); t = 8.0 min, 42/58 (v/v); t = 11.50 min, 42/58 (v/v); t = 11.51 min, 55/45 (v/v); t = 13.00 min, 55/45 (v/v) and 4 µL sample was injected. Peak areas were measured by the integrator and transformed into concentration using the corresponding standard curves. For substrates **2** and **3**, the employed wavelength was 210 and 310 nm respectively, to detect the generated products.

HPLC-ESI-MS for confirmation. To confirm generation of the hydroxylated products **2 a**, **3 a** and **3 b**, HPLC-ESI-MS was employed. An Agilent 6130 Quadrupole LC System (Santa Clara, US), was operated using an identical HPLC method and column as described in section 1.7, above. The ESI-MS was run using the following parameters: API-ES spray chamber with a drying gas flow of 12 L min⁻¹, nebulizer pressure of 50 psi, drying gas temperature 350 °C and a capillary voltage of 3500 V positive and negative. The mass detection for **2 a**, **3 a** and **3 b** was performed in positive SIM-modus with m/z = 150 [M + H], 172 [M + H] and 190 [M + H], respectively.

Chiral HPLC-DAD. Chiral HPLC-DAD measurements for **1 a** were performed according to Wissner and colleagues.[17] Chiral assessments of the products **2 a** and **3 a** were performed on an Agilent 1260 Infinity HPLC-DAD system (Santa Clara, US), equipped with a normal phase Chiralpak IC column (5 µm, 4.6 × 250 mm, Daicel, Osaka, JP), heated to 30 °C. Analysis was performed with a flow rate of 1.4 mL min⁻¹ and 15 µL sample injection. For the separation of the **2 a** enantiomers, n-hexane/isopropanol (95/5) was used as mobile phase and hold isocratically for 22 min. For the separation of **3 a** enantiomers, n-hexane/isopropanol (90/10) was used and hold for 90 min.

Peaks were detected at 210 nm and 310 nm, respectively (Supplementary section 1, Figs. S5 and S12).

NMR for structural characterization and measurement of the optical rotation. The purity of the biosynthesized products was assessed by ^{1}H- and ^{13}C-NMR. Spectra were recorded on a Bruker Avance 500 spectrometer operating at 500.15 and 125.76 MHz, for ^{1}H- and ^{13}C-NMR, respectively. Spectra were recorded using DMSO-d6 for **2 a** and **3 b**, and CDCl$_3$ for **2 b** and **3 a** as solvent. The chemical shifts (δ) are reported in parts per million (ppm) relative to the standard tetramethylsilane (TMS, δ = 0). For identification of **3 a**, COSY-, HSQC- and HMBC-NMRs were performed. Optical rotations were determined using a Perkin Elmer 241 polarimeter at 589 nm and 22 °C.

(+)-(R)-1,2,3,4-Tetrahydroquinoline-4-ol 2 a (106 mg, yellow oil), biosynthesized with TDO variant F114H_A223T: $[\alpha]_D^{22} =$ + 128 (c 0.1, CHCl$_3$), chiral HPLC-DAD of raw product: 94% ee; chiral HPLC-DAD of pure product: 80% ee due to racemization during purification. ^{1}H-NMR NMR (DMSO-d6, 500.15 MHz) δ = 7.08 (d, 1 H), 6.91 (t, 1H), 6.46 (m, 2H), 5.76 (s, 1H), 4.96 (d, 1H), 4.53 (m, 1H), 3.23 (m, 1H), 3.12 (m, 1H), 1.69-1.82 (m, 2H) ppm. ^{13}C-NMR (DMSO-d6, 125.76 MHz) δ = 145.12, 129.34, 127.70, 123.28, 114.69, 113.40, 63.80, 36.18, 30.39 ppm. Results were in agreement with literature.[25,36]

Quinoline 2 b (20 mg, colorless oil), biosynthesized with TDO variant F366V: ^{1}H-NMR (CDCl$_3$, 500.15 MHz) δ = 8.92 (dd, 1 H), 8.13 (dd, 2 H), 7.81 (d, 1 H), 7.72 (ddd, 1 H), 7.54 (ddd, 1 H), 7.39 (dd, 1 H) ppm. ^{13}C-NMR (CDCl$_3$, 125.76 MHz) δ = 150.41, 148.30, 136.07, 129.47, 128.30, 127.80, 126.55, 121.08 ppm. Results were in agreement with literature.[37]

(+)-(1S,2R)-3-(Pyridin-2-yl)cyclohexa-3,5-diene-1,2-diol 3 a (114 mg, yellow oil) biosynthesized with TDO variant M220A_V309G: $[\alpha]_D^{22} =$ + 169 (c 0.1, MeOH), chiral HPLC-DAD: > 98% ee. ^{1}H-NMR (CDCl$_3$, 500.15 MHz) δ = 8.53 (d, 1H), 7.71 (dd, 1H), 7.63 (d, 1H), 7.19 (dd, 1H), 6.66 (d, 1H), 6.14-6.22 (m, 2H), 4.92 (d, 1H), 4.50 (m, 1H) ppm. ^{13}C-NMR (CDCl$_3$, 125.76 MHz) δ = 157.15, 148.01, 137.14, 136.12, 131.75, 124.69, 124.88, 122.16, 120.08, 69.04, 67.41 ppm. Results were in agreement with literature.[11]

2-Phenylpyridin-3-ol 3 b (6 mg, white solid) biosynthesized with TDO variant M220A_V309G: ^{1}H-NMR NMR (DMSO-d6, 500.15 MHz) δ = 10.14 (s, 1H), 8.17 (d, 1H), 8.01 (d, 2H), 7.44 (t, 2H), 7.35 (t, 2H), 7.21 (dd, 1H) ppm. ^{13}C-NMR (DMSO-d6, 125.76 MHz) δ = 151.48, 144.38, 140.19, 137.99, 128.79, 127.68, 123.49, 123.31 ppm. Results were in agreement with literature.[38]

Acknowledgements

References

[1] M. A. Endoma, V. P. Bui, J. Hansen, T. Hudlicky, *Org. Process Res. Dev.* **2002**, *6*, 525–532.

[2] T. Hudlicky, *ACS Omega* **2018**, *3*, 17326–17340.

[3] D. R. Boyd, N. D. Sharma, C. J. McGivern, P. J. Stevenson, P. Hoering, C. C. R. Allen, *J. Org. Chem.* **2019**, *84*, 15165–15172.

[4] F. F. Özgen, S. Schmidt, *Biocatalysis* **2019**, 57–82.

[5] S. Resnick, K. Lee, D. Gibson, *J. Ind. Microbiol. Biotechnol.* **1996**, *17*, 438–457.

[6] D. R. Boyd, N. D. Sharma, S. A. Haughey, M. A. Kennedy, B. T. McMurray, G. N. Sheldrake, C. C. R. Allen, H. Dalton, K. Sproule, *J. Chem. Soc. Perkin Trans. 1* **1998**, 1929–1934.

[7] D. S. Torok, S. M. Resnick, J. M. Brand, D. L. Cruden, D. T. Gibson, *J. Bacteriol.* **1995**, *177*, 5799–5805.

[8] M. A. Vila, V. Steck, S. Rodriguez Giordano, I. Carrera, R. Fasan, *ChemBioChem* **2020**, *21*, 1981–1987.

[9] R. E. Parales, S. M. Resnick, *Biocatalysis in the Pharmaceutical and Biotechnology Industries*, CRC Press, **2006**, 299–331.

[10] M. A. Vila, M. Brovetto, D. Gamenara, P. Bracco, G. Zinola, G. Seoane, S. Rodríguez, I. Carrera, *J. Mol. Catal. B* **2013**, *96*, 14–20.

[11] D. R. Boyd, N. D. Sharma, G. P. Coen, F. Hempenstall, V. Ljubez, J. F. Malone, C. C. R. Allen, J. T. G. Hamilton, *Org. Biomol. Chem.* **2008**, *6*, 3957–3966.

[12] T. Sakamoto, J. M. Joern, A. Arisawa, F. H. Arnold, *Appl. Environ. Microbiol.* **2001**, *67*, 3882–3887.

[13] L. M. Newman, H. Garcia, T. Hudlicky, S. A. Selifonov, *Tetrahedron* **2004**, *60*, 729–734.

[14] N. Zhang, B. G. Stewart, J. C. Moore, R. L. Greasham, D. K. Robinson, B. C. Buckland, C. Lee, *Metab. Eng.* **2000**, *2*, 339–348.

[15] M. A. Vila, D. Umpiérrez, G. Seoane, S. Rodríguez, I. Carrera, N. Veiga, *J. Mol. Catal. B* **2016**, *133*, S410–S419.

[16] D. Boyd, N. Sharma, I. Brannigan, C. McGivern, P. Nockemann, P. Stevenson, C. McRoberts, P. Hoering, C. Allen, *Adv. Synth. Catal.* **2019**, *361*, 2526–2537.

[17] J. L. Wissner, W. Escobedo-Hinojosa, A. Vogel, B. Hauer, *J. Biotechnol.* **2021**, *326*, 37–39.

[18] M. A. Vila, D. Umpiérrez, N. Veiga, G. Seoane, I. Carrera, S. Rodríguez Giordano, *Adv. Synth. Catal.* **2017**, *359*, 2149–2157.

[19] R. Friemann, K. Lee, E. N. Brown, D. T. Gibson, H. Eklund, S. Ramaswamy, *Acta Crystallogr. Sect. D* **2009**, *65*, 24–33.

[20] B. Kauppi, K. Lee, E. Carredano, R. E. Parales, D. T. Gibson, H. Eklund, S. Ramaswamy, *Structure* **1998**, *6*, 571–586.

Adv. Synth. Catal. **2021**, *363*, 1–11   **Wiley Online Library**   9

These are not the final page numbers! ↗↗

© 2021 The Authors. Advanced Synthesis & Catalysis published by Wiley-VCH GmbH

126

[21] M. Mohammadi, J. F. Viger, P. Kumar, D. Barriault, J. T. Bolin, M. Sylvestre, *J. Biol. Chem.* **2011**, *286*, 27612–27621.

[22] M. Nettekoven, J. Fingerle, U. Grether, S. Grüner, A. Kimbara, B. Püllmann, M. Rogers-Evans, S. Röver, F. Schuler, T. Schulz-Gasch, C. Ullmer, *Bioorg. Med. Chem. Lett.* **2013**, *23*, 1177–1181.

[23] M. Okamoto, O. Sakuragi, Y. Mori, M. Kishida, T. Higashijima, *Optically Active Cyclic Alcohol Compound*, **2013**, Patent No. US8471028B2.

[24] V. Sridharan, P. A. Suryavanshi, J. C. Menéndez, *Chem. Rev.* **2011**, *111*, 7157–7259.

[25] D. Zheng, M. Yang, J. Zhuo, K. Li, H. Zhang, J. Yang, B. Cui, Y. Chen, *J. Mol. Catal. B* **2014**, *110*, 87–91.

[26] D. R. Boyd, G. N. Sheldrake, *Nat. Prod. Rep.* **1998**, *15*, 309–324.

[27] S. Beil, J. R. Mason, K. N. Timmis, D. H. Pieper, *J. Bacteriol.* **1998**, *180*, 5520–5528.

[28] J. M. Halder, B. M. Nestl, B. Hauer, *ChemCatChem* **2018**, *10*, 178–182.

[29] C. Gally, B. M. Nestl, B. Hauer, *Angew. Chem. Int. Ed.* **2015**, *54*, 12952–12956; *Angew. Chem.* **2015**, *127*, 13144–13148.

[30] E. L. Ang, J. P. Obbard, H. Zhao, *FEBS J.* **2007**, *274*, 928–939.

[31] R. E. Parales, K. Lee, S. M. Resnick, H. Jiang, D. J. Lessner, D. T. Gibson, *J. Bacteriol.* **2000**, *182*, 1641–1649.

[32] R. E. Parales, S. M. Resnick, C. L. Yu, D. R. Boyd, N. D. Sharma, D. T. Gibson, *J. Bacteriol.* **2000**, *182*, 5495–5504.

[33] H. Suenaga, T. Watanabe, M. Sato, Ngadiman, K. Furukawa, *J. Bacteriol.* **2002**, *184*, 3682–3688.

[34] J. L. Wissner, J. Ludwig, W. Escobedo-Hinojosa, B. Hauer, *J. Biotechnol.* **2020**, *325*, 380–388.

[35] V. J. Kolcsár, F. Fülöp, G. Szőllősi, *ChemCatChem* **2019**, *11*, 2725–2731.

[36] C. Yin, X. Q. Dong, X. Zhang, *Adv. Synth. Catal.* **2018**, *360*, 4319–4324.

[37] Y. Fukazawa, A. E. Rubtsov, A. V. Malkov, *Eur. J. Org. Chem.* **2020**, *2020*, 3317–3319.

[38] W. X. Huang, B. Wu, X. Gao, M. W. Chen, B. Wang, Y. G. Zhou, *Org. Lett.* **2015**, *17*, 1640–1643.

These are not the final page numbers! ↗↗

6.6 RESEARCH ARTICLE III

Supplementary Information

Supporting Information

Semi-Rational Engineering of Toluene Dioxygenase from *Pseudomonas putida* F1 towards Oxyfunctionalization of Bicyclic Aromatics

Julian L. Wissner, Jona T. Schelle, Wendy Escobedo-Hinojosa, Andreas Vogel, and Bernhard Hauer,*

Supplementary Data

Semi-Rational Engineering of Toluene Dioxygenase from *Pseudomonas putida* F1 towards Oxyfunctionalization of Bicyclic Aromatics

Julian L. Wissner[a], Jona T. Schelle[a], Wendy Escobedo-Hinojosa[a], Andreas Vogel[b], Bernhard Hauer[a]*

[a]Institute of Technical Biochemistry, University of Stuttgart, Allmandring 31, 70569 Stuttgart, Germany

[b]c-LEcta GmbH, Perlickstr. 5, 04103 Leipzig, Germany

*Corresponding author

E-mail address: bernhard.hauer@itb.uni-stuttgart.de

Table of Contents

Supplementary section 1: HPLC-Chromatograms

Fig. S1: HPLC-DAD chromatogram (210 nm) of the commercial standards for the expected products 1,2,3,4-tetrahydroquinolin-4-ol **2a** (green; 2.07 min; racemic), quinoline **2b** (red; 4.38 min), and the substrate 1,2,3,4-tetrahydroquinoline **2** (blue; 5.82 min).
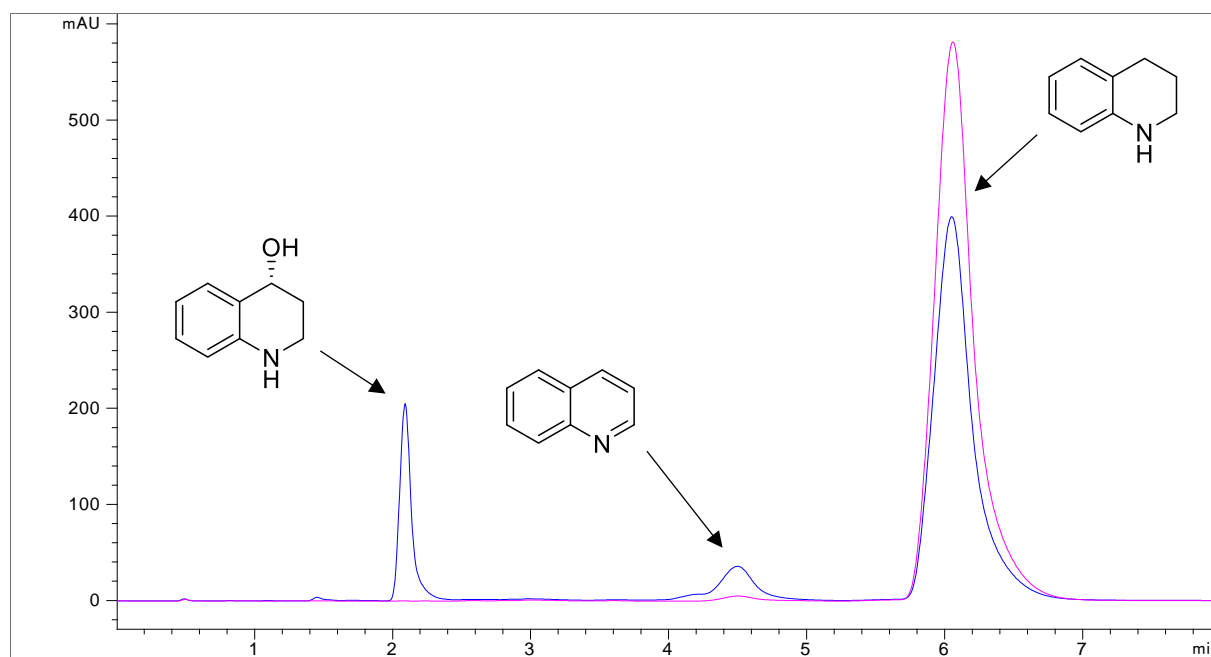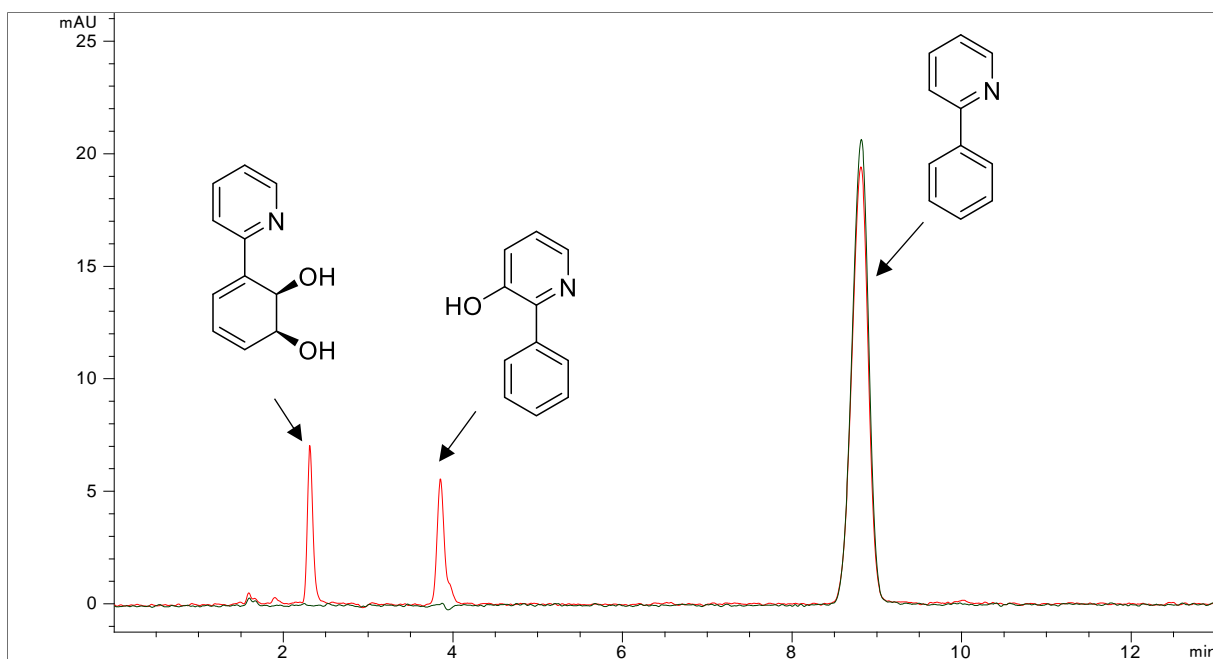

Fig. S2: HPLC-DAD chromatogram (210 nm) of the 1,2,3,4-tetrahydroquinoline **2** biotransformation, catalyzed by *E. coli* BW25113, harboring either pBAD18-TDO (blue) or pBAD18 empty vector (pink) in 0.1 mM potassium phosphate buffer pH 7.4. Generated products (*R*)-1,2,3,4-tetrahydroquinolin-4-ol **2a** (2.07 min), quinoline **2b** (4.38 min), and substrate 1,2,3,4-tetrahydroquinoline **2** (5.82 min).
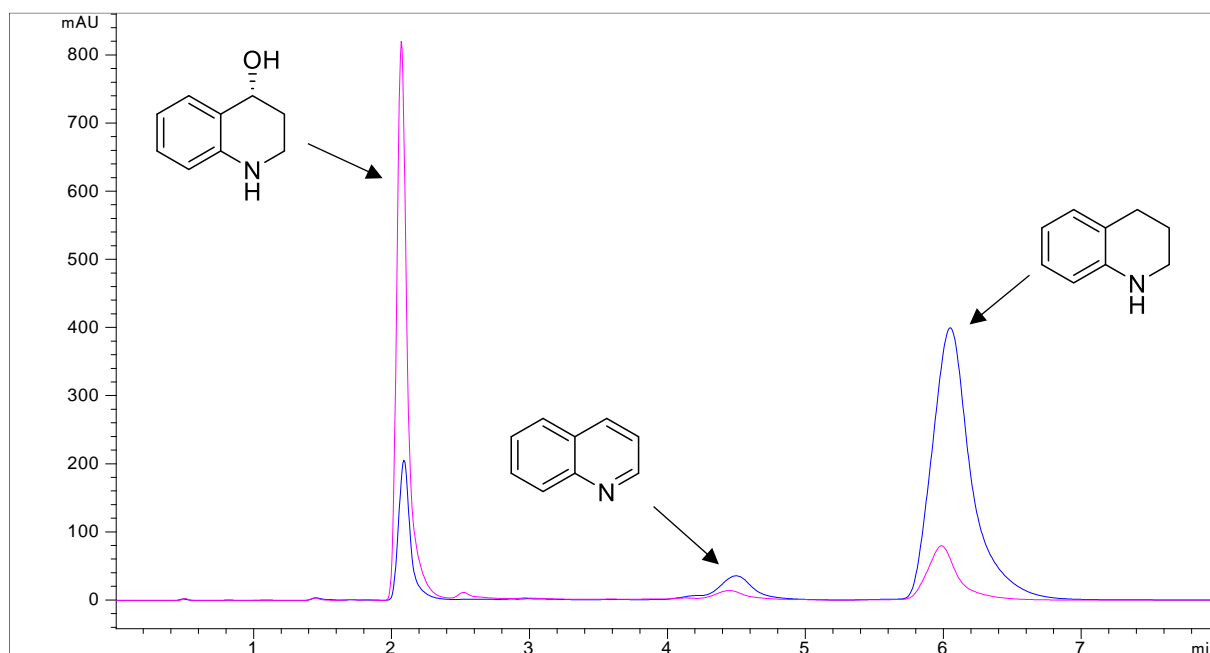
Fig. S3: HPLC-DAD chromatogram (210 nm) of the 1,2,3,4-tetrahydroquinoline **2** biotransformation, catalyzed either by TDO wild type (blue) or by variant TDO$_{F114H\_A223T}$ (pink) in 0.1 mM potassium phosphate buffer pH 7.4. Generated products (*R*)-1,2,3,4-tetrahydroquinolin-4-ol **2a** (2.07 min), quinoline **2b** (4.38 min), and substrate 1,2,3,4-tetrahydroquinoline **2** (5.82 min).
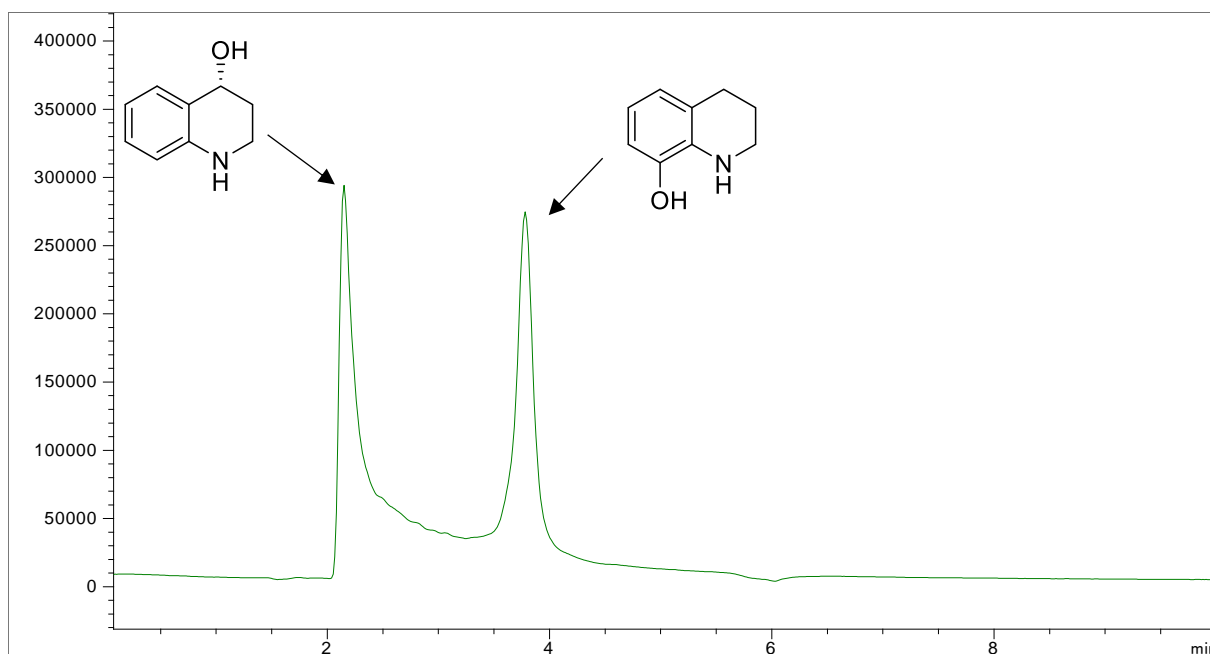
Fig. S4: HPLC-ESI-MS chromatogram of the 1,2,3,4-tetrahydroquinoline **2** biotransformation, catalyzed by TDO wild type in 0.1 mM potassium phosphate buffer pH 7.4. Samples were measured in SIM mode for monohydroxylated products with m/z =150 [M+H]. Generated product (*R*)-1,2,3,4-tetrahydroquinolin-4-ol **2a** (2.07 min). Minor side product corresponding to the peak recorded at 3.76 min was identified as 1,2,3,4-tetrahydroquinolin-8-ol by employing a commercial standard. 1,2,3,4-tetrahydroquinolin-8-ol was not further considered, since its concentration was as low as <0.05 mM.
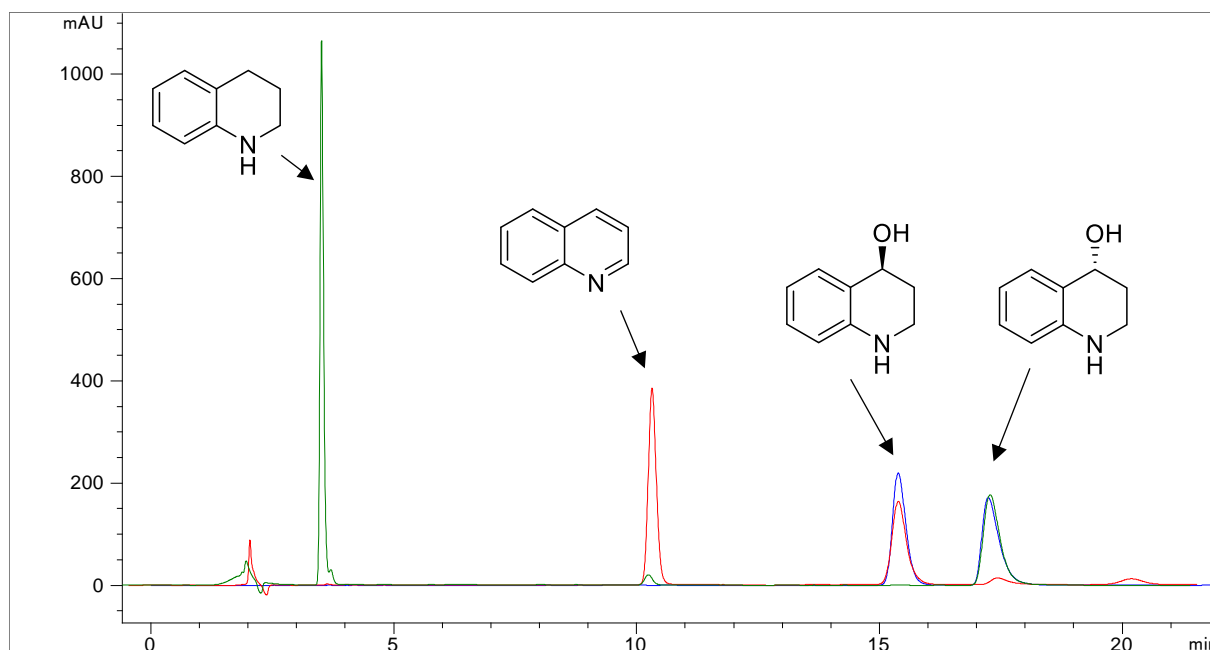
Fig. S5: Chiral HPLC-DAD (210 nm) chromatogram of the racemic commercial standard 1,2,3,4-tetrahydroquinolin-4-ol (blue), chemically synthesized (*S*)-1,2,3,4-tetrahydroquinolin-4-ol (red), and biosynthesized (*R*)-1,2,3,4-tetrahydroquinolin-4-ol **2a** employing TDO wild type (green). Recorded peaks; 1,2,3,4-tetrahydroquinoline **2** (3.51 min), quinoline **2b** (10.18 min), (*S*)-1,2,3,4-tetrahydroquinolin-4-ol (15.28 min), and (*R*)-1,2,3,4-tetrahydroquinolin-4-ol **2a** (17.24 min).
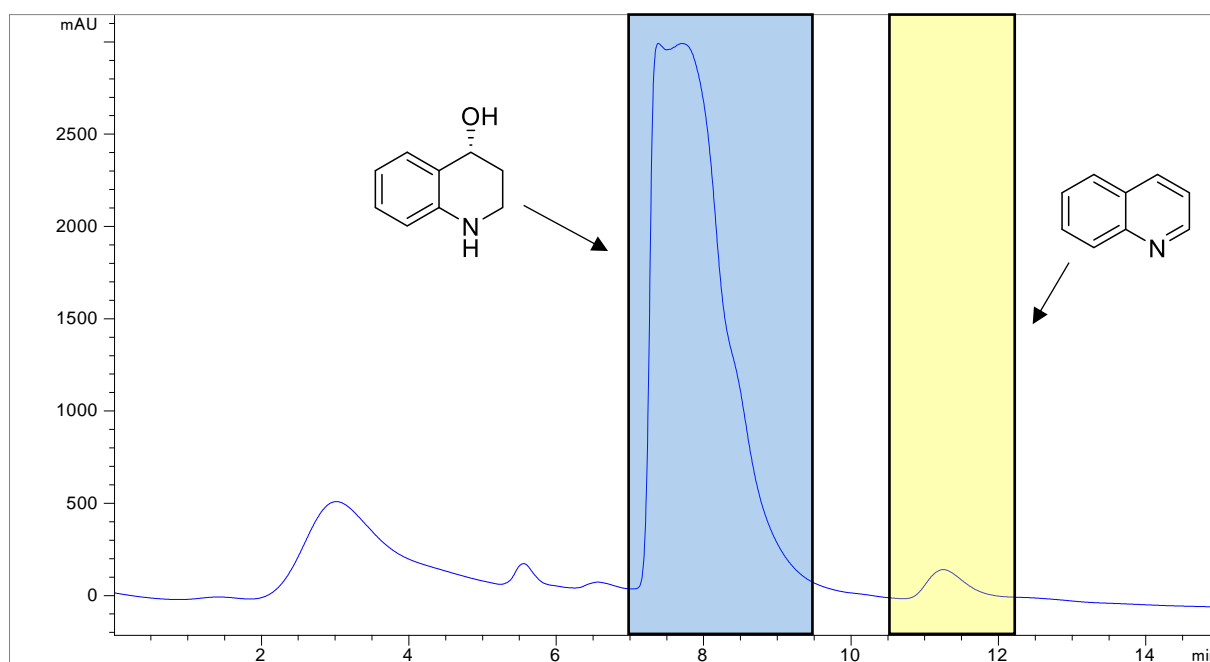
Fig. S6: HPLC-DAD (210 nm) chromatogram of the preparative HPLC purification method for products of the 1,2,3,4-tetrahydroquinoline **2** biotransformation, catalyzed by TDO_{F114H_A223T} in 0.1 mM TRIS buffer pH 8.5. (*R*)-1,2,3,4-Tetrahydroquinolin-4-ol **2a** was collected from minute 7.03 to 8.49 (light blue area) and quinoline **2b** from minute 10.50 to 12.27 (yellow area).
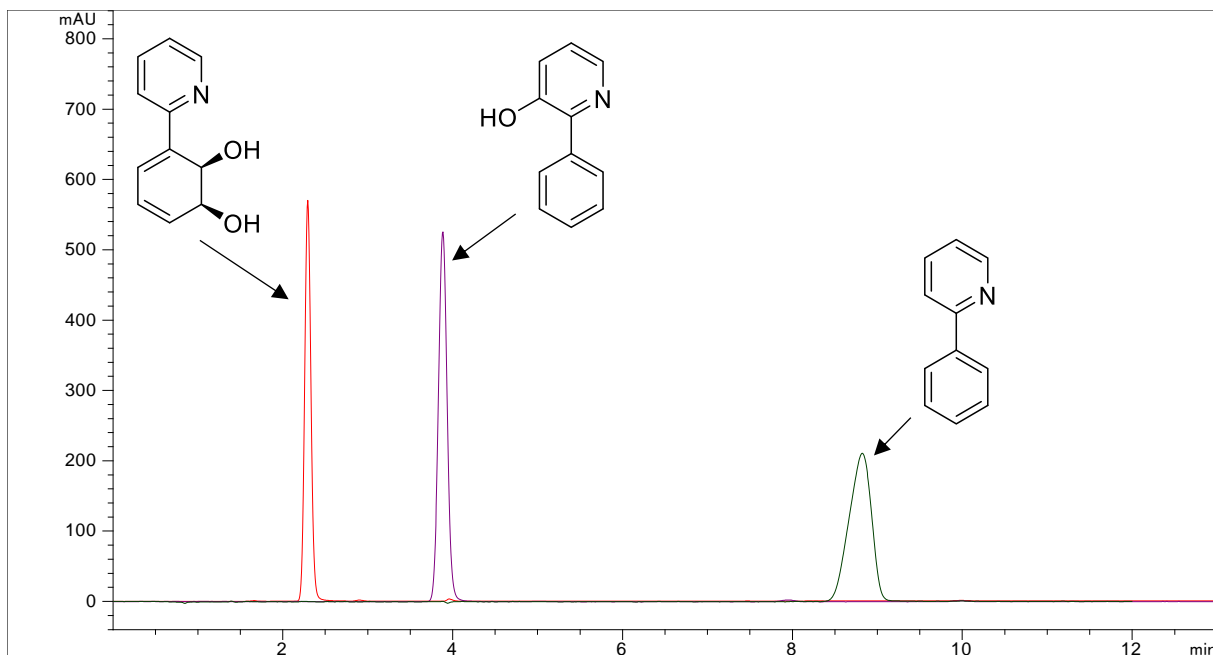
Fig. S7: HPLC-DAD chromatogram (310 nm) of the biosynthesized product standard (+)-(1*S*,2*R*)-3-(pyridin-2-yl)cyclohexa-3,5-diene-1,2-diol **3a** (red; 2.34 min). Commercial standard peaks of the product 2-phenylpyridin-3-ol **3b** and the substrate 2-phenylpyridine **3**, are shown in (purple; 3.87 min) and (green; 8.95 min), respectively.
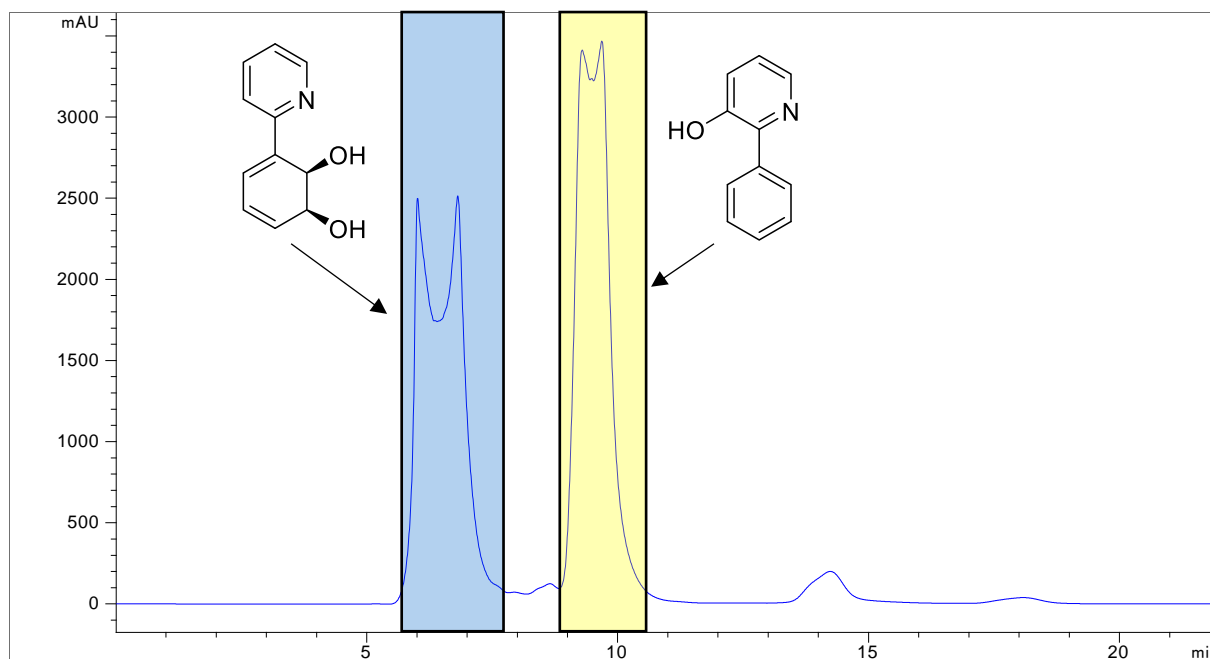

Fig. S8: HPLC-DAD chromatogram (310 nm) of the 2-phenylpyridine **3** biotransformation, catalyzed by *E. coli* BW25113, harboring either pBAD18-TDO (red) or pBAD18 empty vector (green). Generated products (+)-(1*S*,2*R*)-3-(pyridin-2-yl)cyclohexa-3,5-diene-1,2-diol **3a** (2.34 min), 2-phenylpyridin-3-ol **3b** (3.87 min) and substrate 2-phenylpyridine **3** (8.95 min).
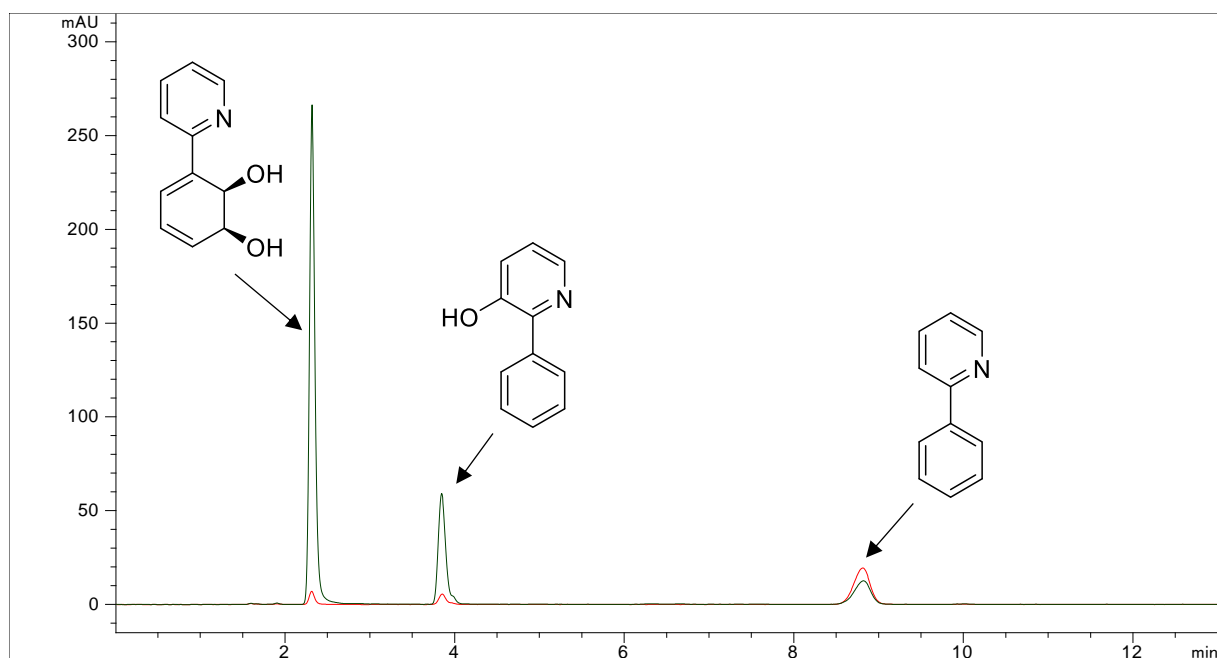
Fig. S9: HPLC-DAD chromatogram (310 nm) of the 2-phenylpyridine **3** biotransformation, catalyzed either by TDO wild type (red), or by variant TDO$_{M220A\_V309G}$ (green). Generated products (+)-(1*S*,2*R*)-3-(pyridin-2-yl)cyclohexa-3,5-diene-1,2-diol **3a** (2.34 min), 2-phenylpyridin-3-ol **3b** (3.87 min) and substrate 2-phenylpyridine **3** (8.95 min).
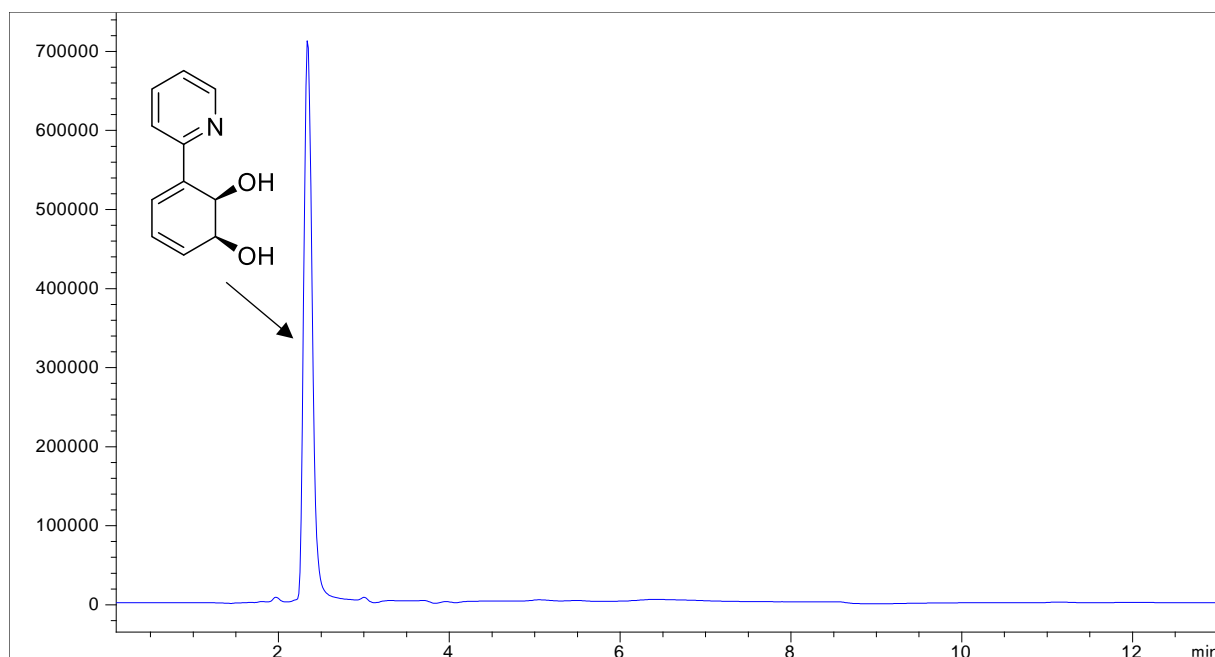
Fig. S10: HPLC-ESI-MS chromatogram of the 2-phenylpyridine **3** biotransformation catalyzed by TDO wild type, measured in SIM mode for dihydroxylated products with m/z =190 [M+H]. Recorded peak corresponded to (+)-(1*S*,2*R*)-3-(pyridin-2-yl)cyclohexa-3,5-diene-1,2-diol **3a** (2.34 min).
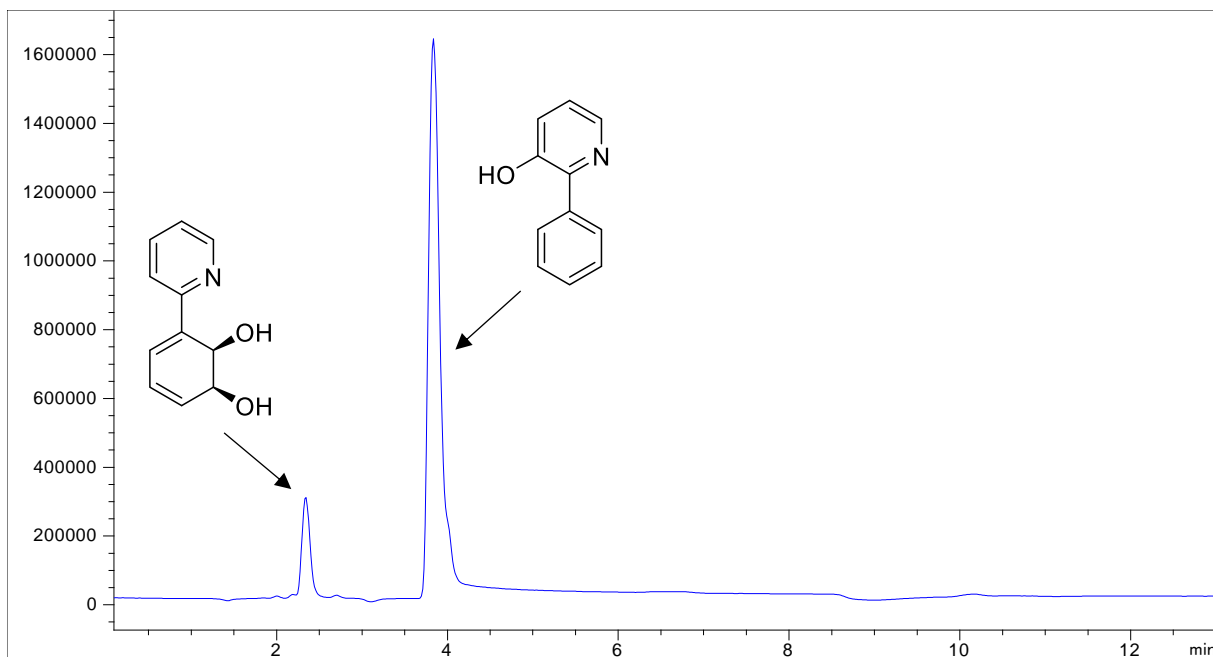
Fig. S11: HPLC-ESI-MS chromatogram of the 2-phenylpyridine **3** biotransformation catalyzed by TDO wild type, measured in SIM mode for monohydroxylated products with m/z =172 [M+H]. Recorded peaks; partially dehydrated (+)-(1*S*,2*R*)-3-(pyridin-2-yl)cyclohexa-3,5-diene-1,2-diol **3a** (2.34 min) and 2-phenylpyridin-3-ol **3b** (3.87 min).
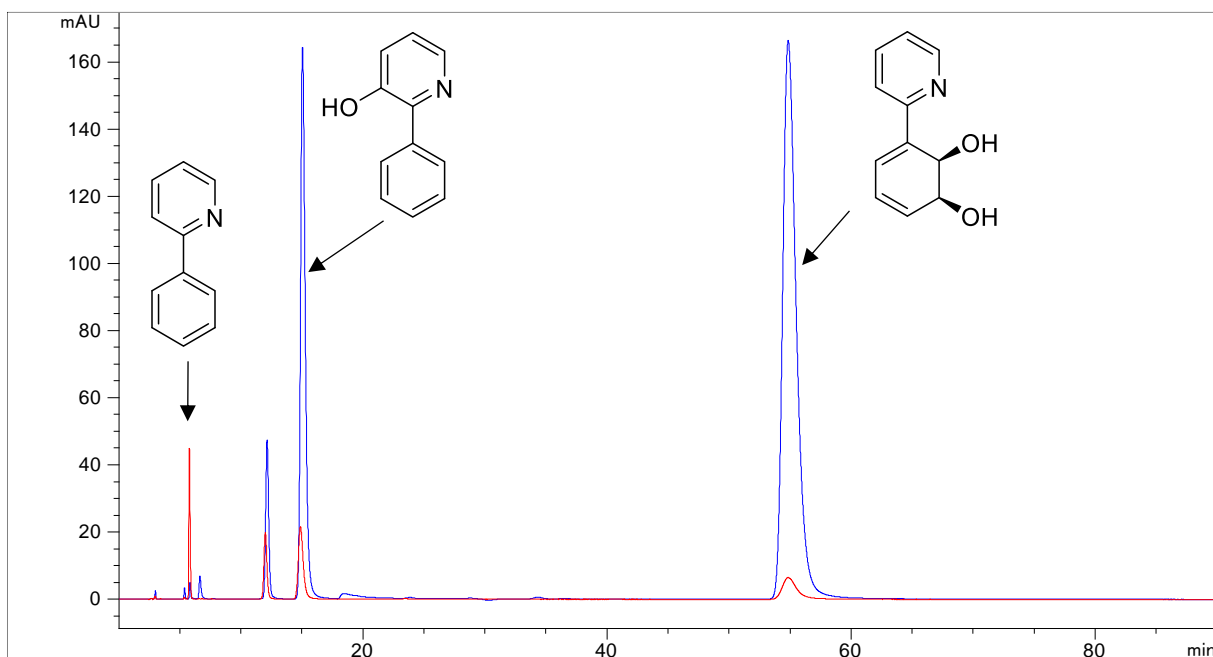
Fig. S12: Chiral HPLC-DAD (310 nm) chromatogram of the 2-phenylpyridine **3** biotransformation catalyzed either by TDO wild type (red), or variant TDO_M220A_V309G (blue). Recorded peaks; 2-Phenypyridine **3** (5.84 min), 2-phenylpyridin-3-ol **3b** (15.06 min), and (+)-(1*S*,2*R*)-3-(pyridin-2-yl)cyclohexa-3,5-diene-1,2-diol **3a** (54.84 min). The peak for enantiomer (-)-(1*R*,2*S*)-3-(pyridin-2-yl)cyclohexa-3,5-diene-1,2-diol could not be identified, due to the exclusive TDO_M220A_V309G catalyzed production of (+)-(1*S*,2*R*)-3-(pyridin-2-yl)cyclohexa-3,5-diene-1,2-diol **3a** (>99% *ee*), verified *via* polarimetry of the isolated product.

Fig. S13: HPLC-DAD (310 nm) chromatogram of the preparative HPLC purification method for products of the 2-phenylpyridine **3** biotransformation, catalyzed by TDO_{F366V}. Recorded peaks; (+)-(1*S*,2*R*)-3-(pyridin-2-yl)cyclohexa-3,5-diene-1,2-diol **3a** was collected from minute 5.76 to 7.83 (light blue area) and 2-phenylpyridin-3-ol **3b** from minute 8.92 to 10.60 (yellow area).

138

9

Supplementary section 2: ^{1}H- and ^{13}C-NMR-spectra
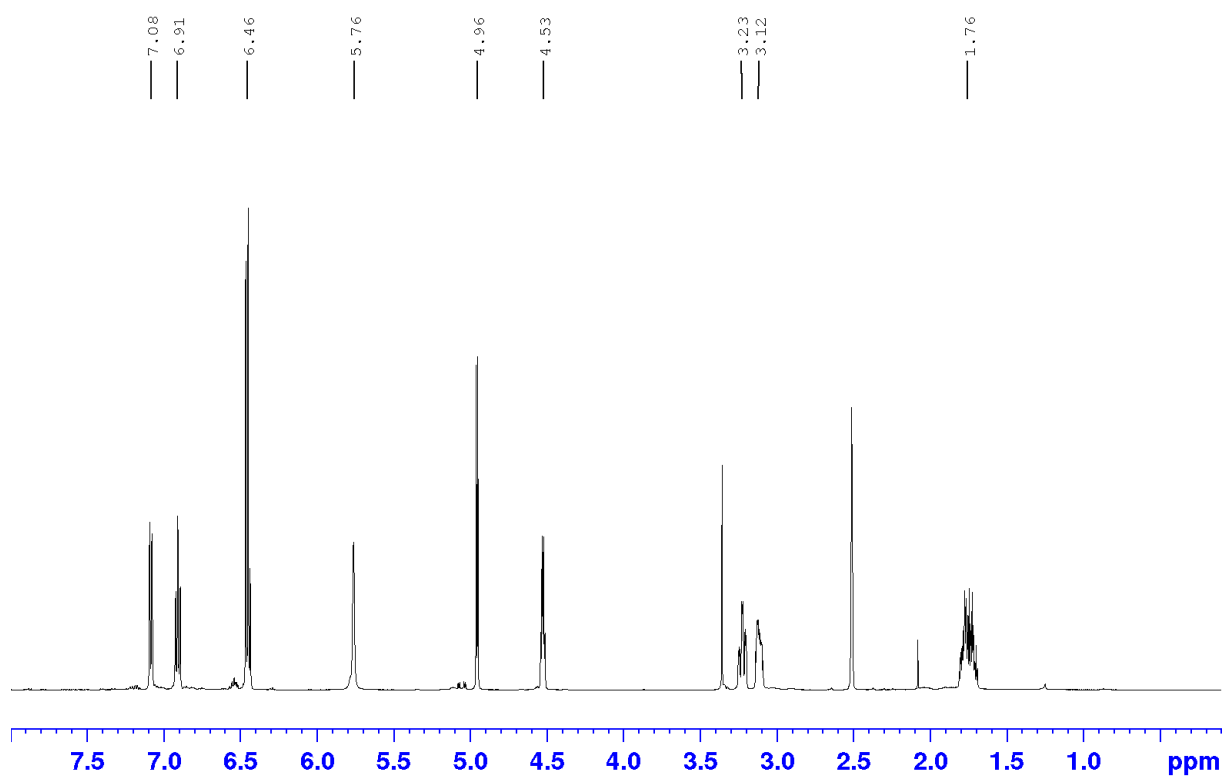
Fig. S14: ^{1}H-NMR of (*R*)-tetrahydroquinolin-4-ol **2a** in DMSO-d6. H_2O and DMSO can be seen at 3.33 and 2.50 ppm, respectively.
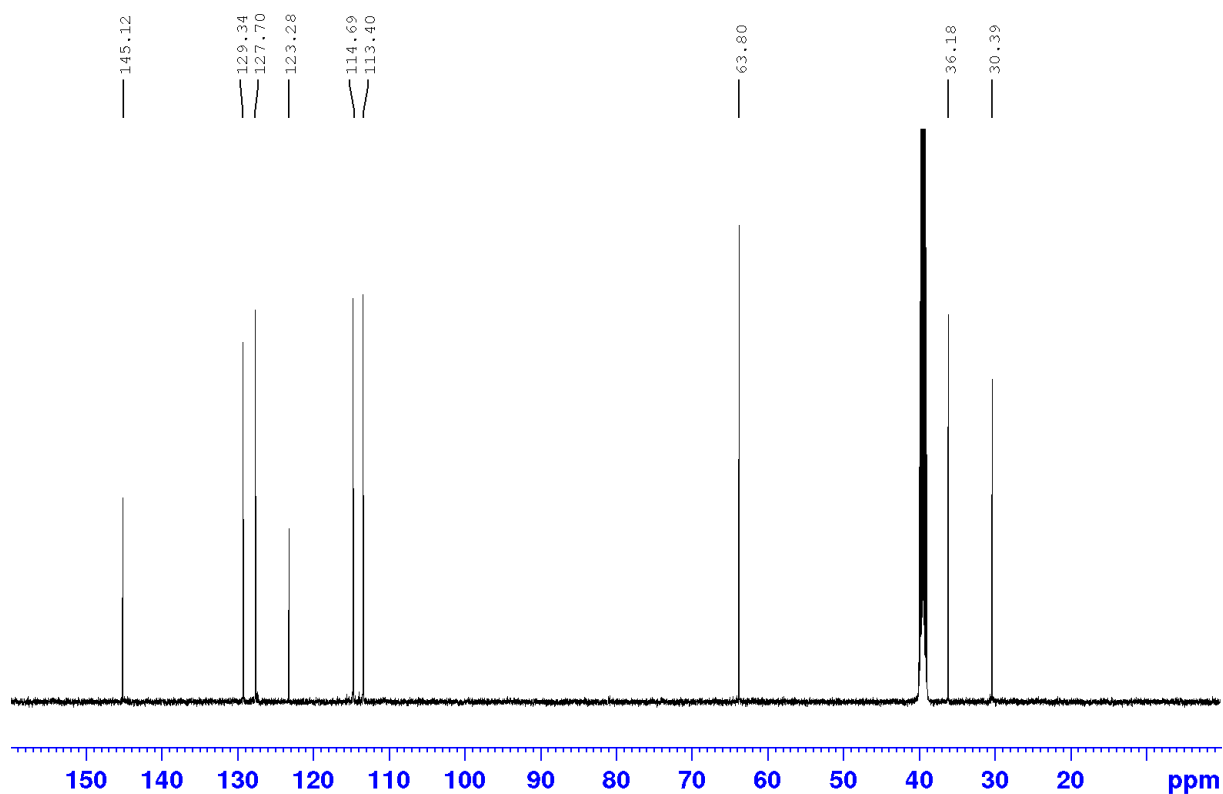
Fig. S15: ^{13}C-NMR of (*R*)-tetrahydroquinolin-4-ol **2a** in DMSO-d6. DMSO can be seen at 39.50 ppm.
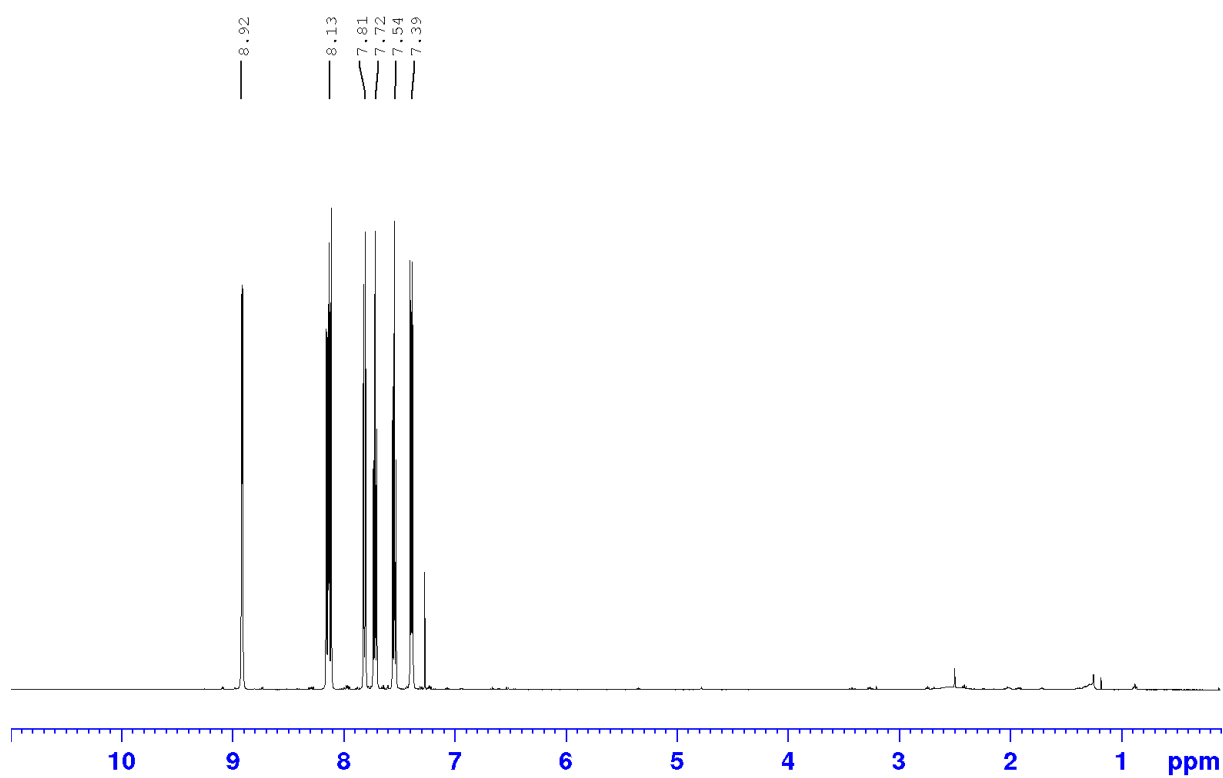
Fig. S16: ^{1}H-NMR of quinoline **2b** in CDCl$_3$. CHCl$_3$ can be seen at 7.27 ppm.
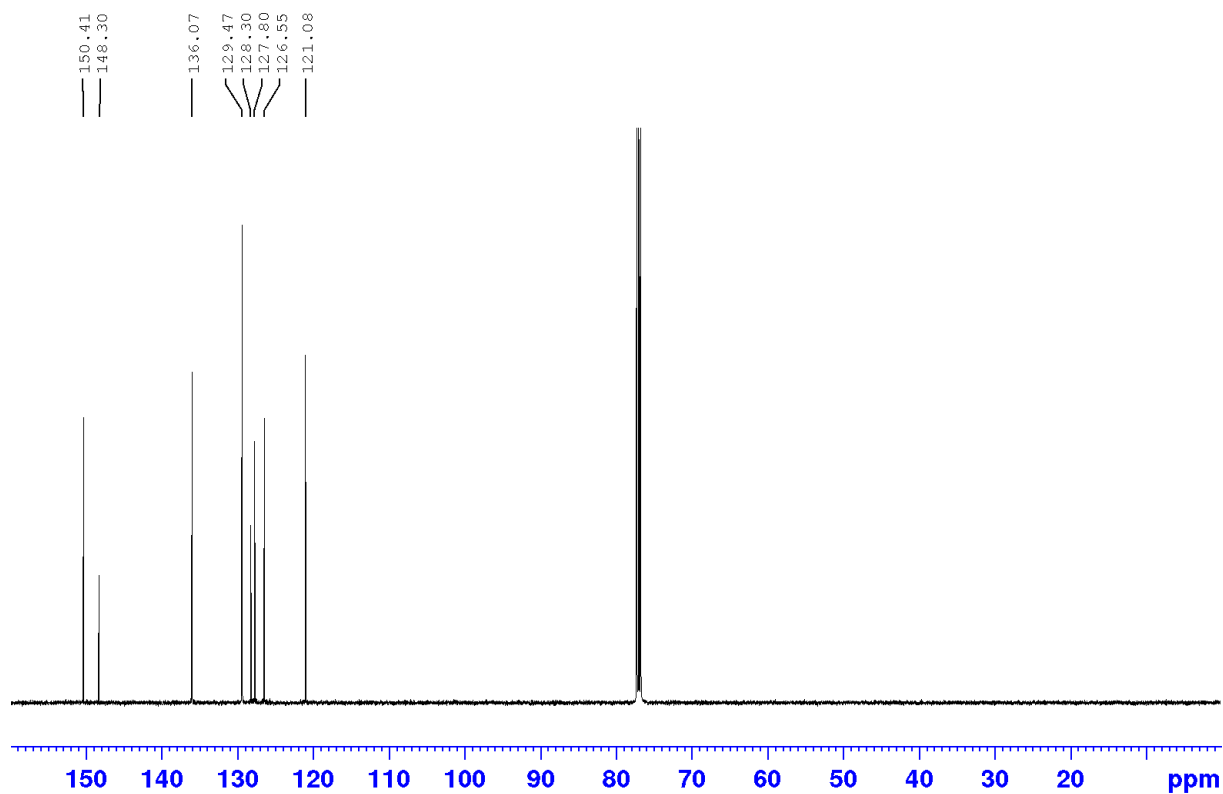
Fig. S17: ^{13}C-NMR of quinoline **2b** in CDCl$_3$. CHCl$_3$ can be seen at 77.16 ppm.
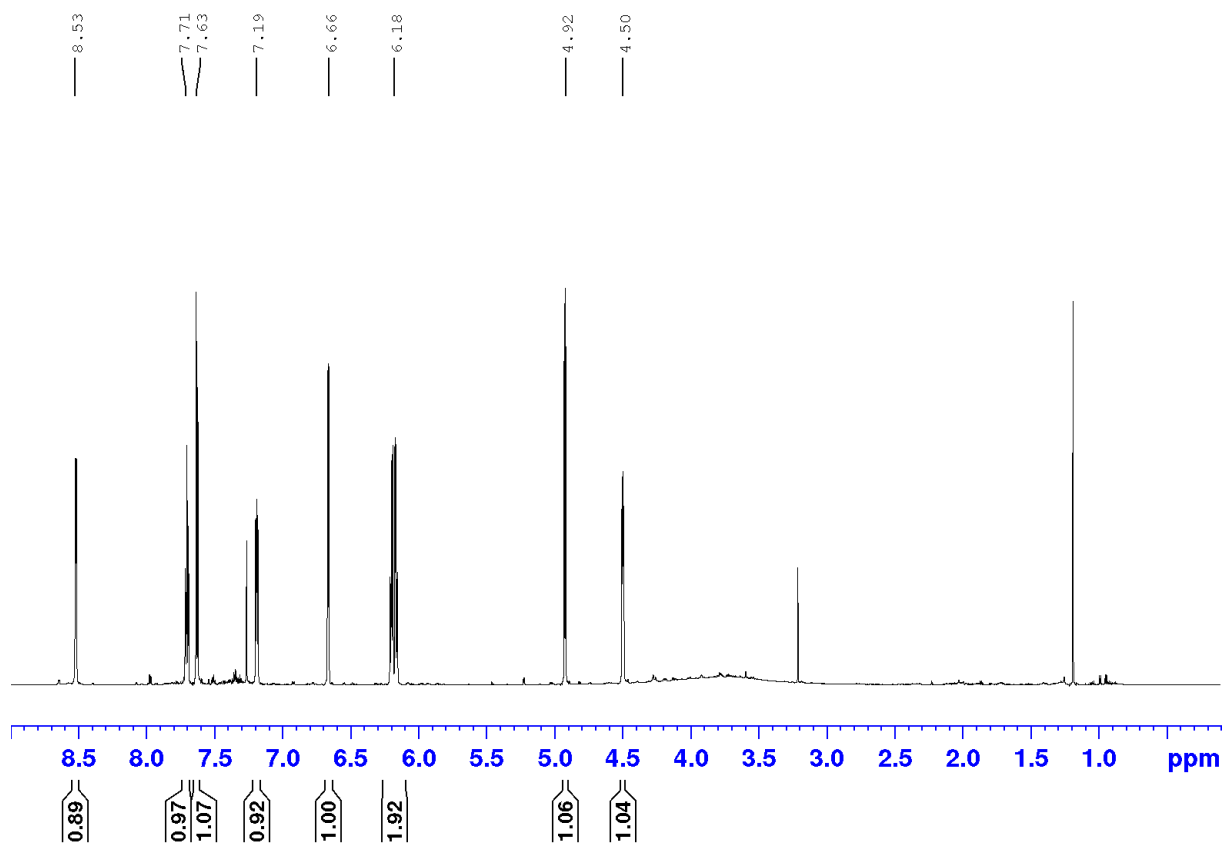
Fig. S18: ^{1}H-NMR of (+)-(1*S*,2*R*)-3-(pyridin-2-yl)cyclohexa-3,5-diene-1,2-diol **3a** in CDCl$_3$. CHCl$_3$ can be seen at 7.27 ppm.
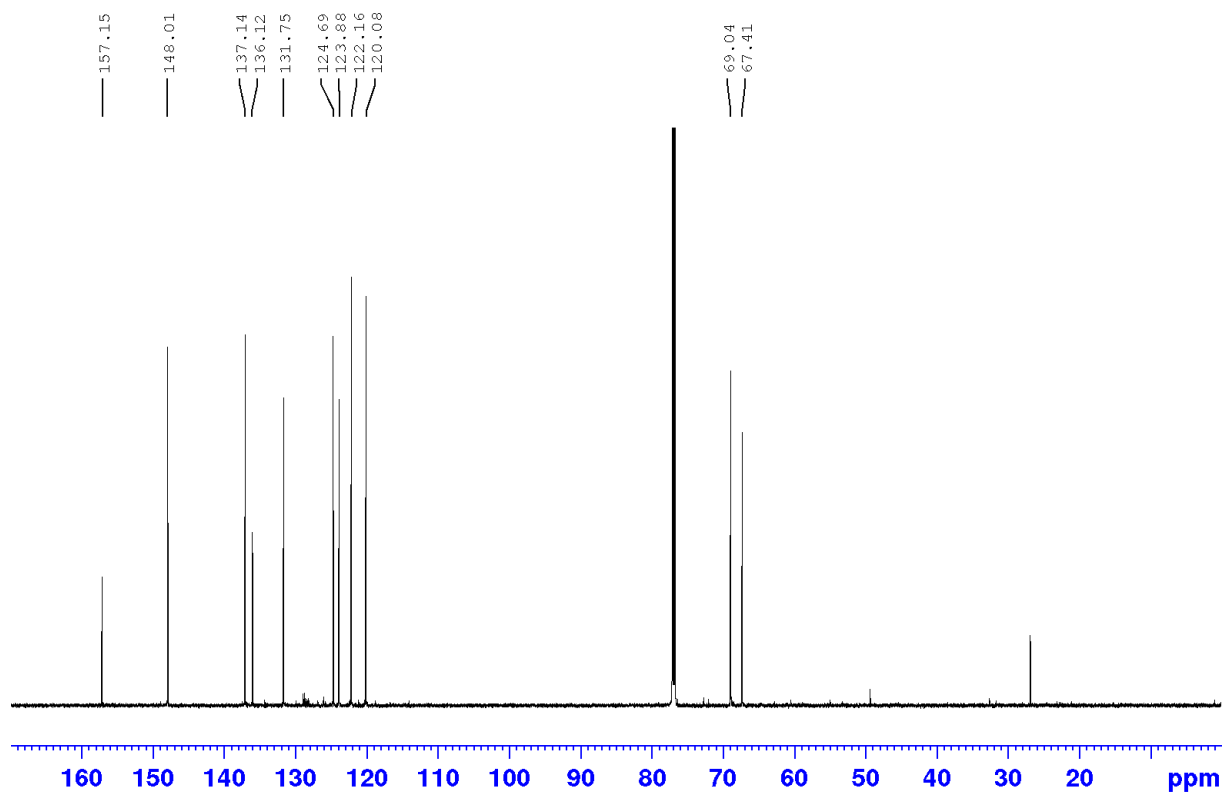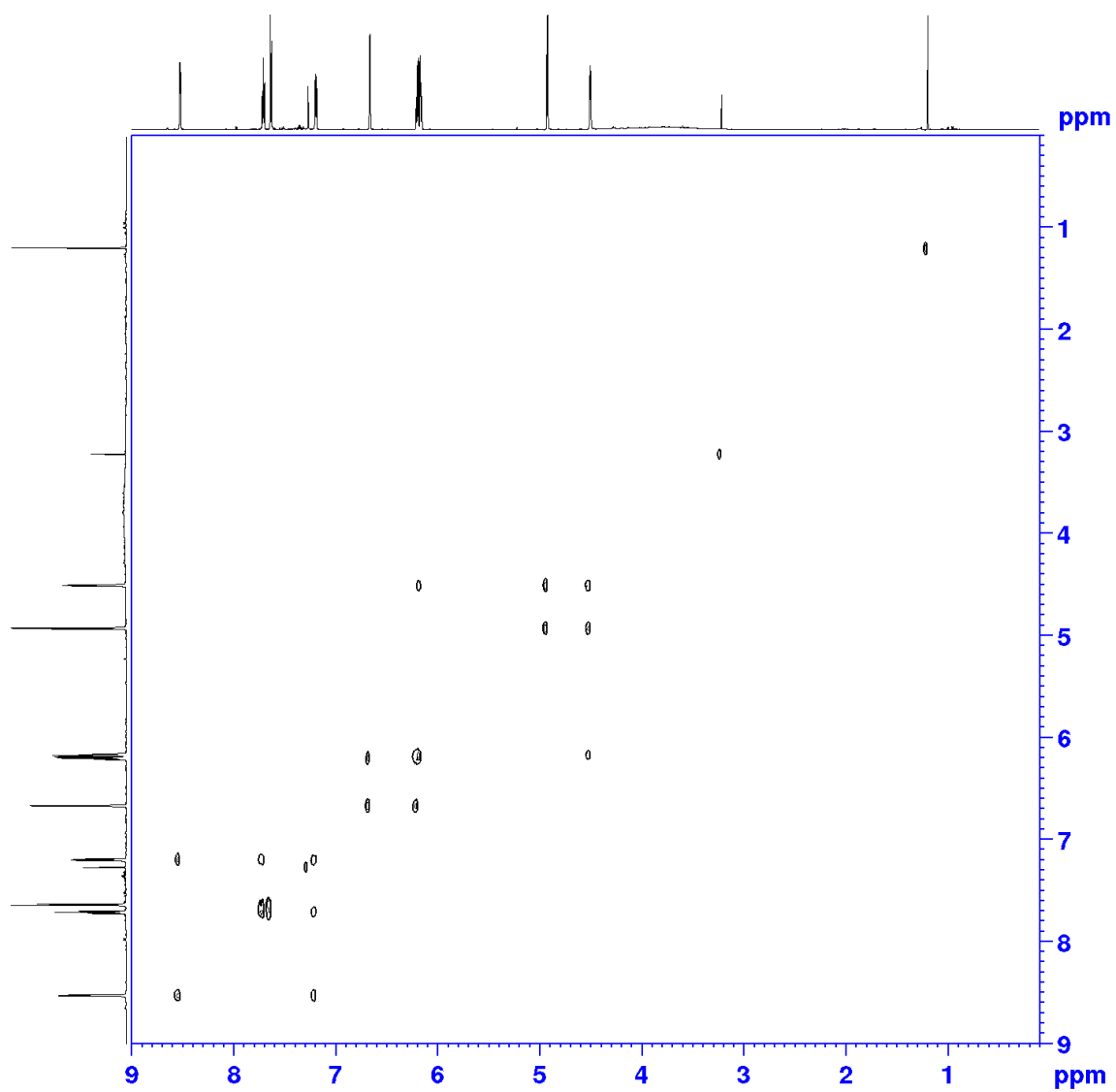

Fig. S19: ^{13}C-NMR of (+)-(1*S*,2*R*)-3-(pyridin-2-yl)cyclohexa-3,5-diene-1,2-diol **3a** in CDCl$_3$. CHCl$_3$ can be seen at 77.16 ppm.
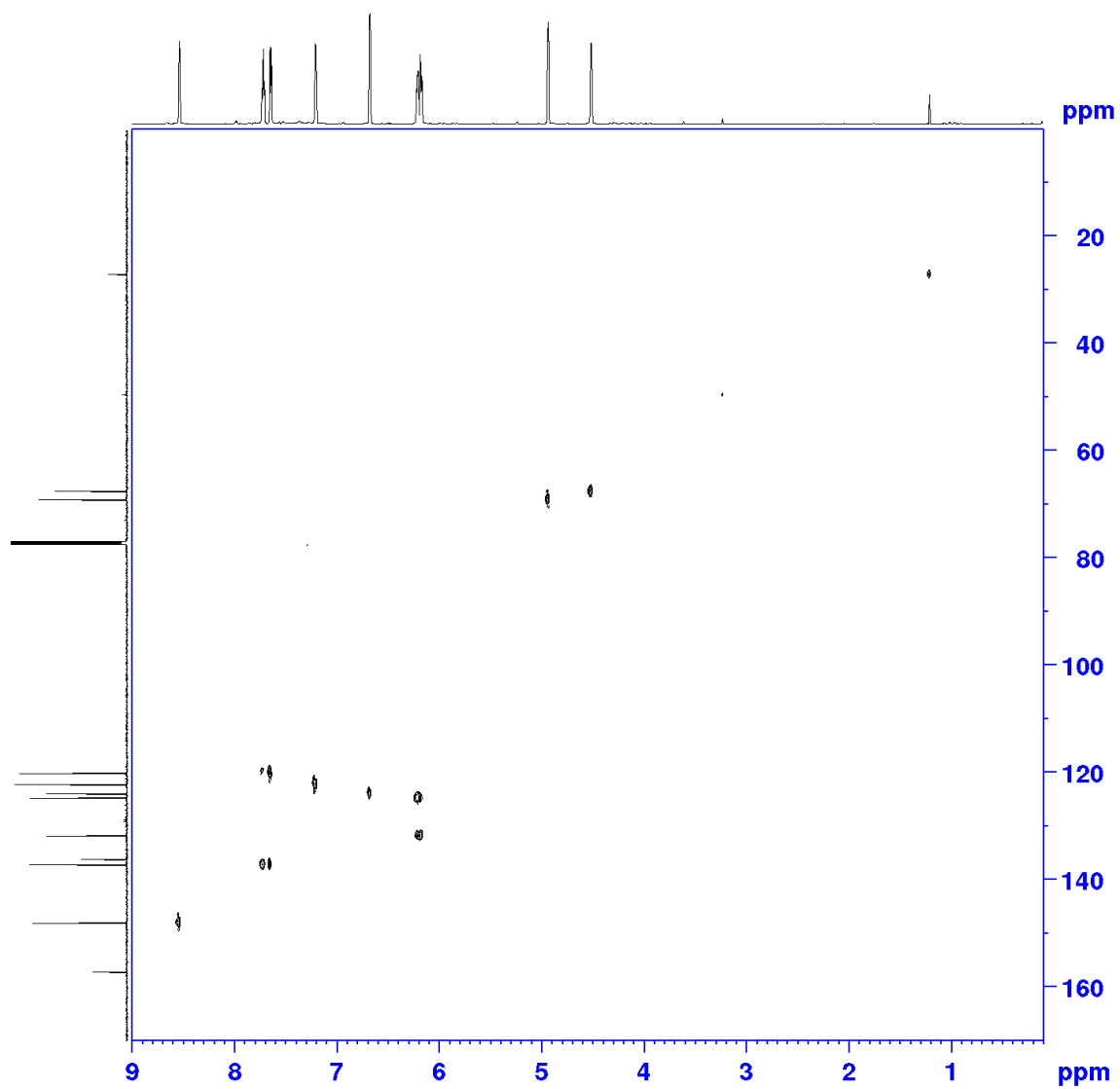
12

Fig. S20: COSY-NMR of (+)-(1*S*,2*R*)-3-(pyridin-2-yl)cyclohexa-3,5-diene-1,2-diol **3a** in CDCl₃.

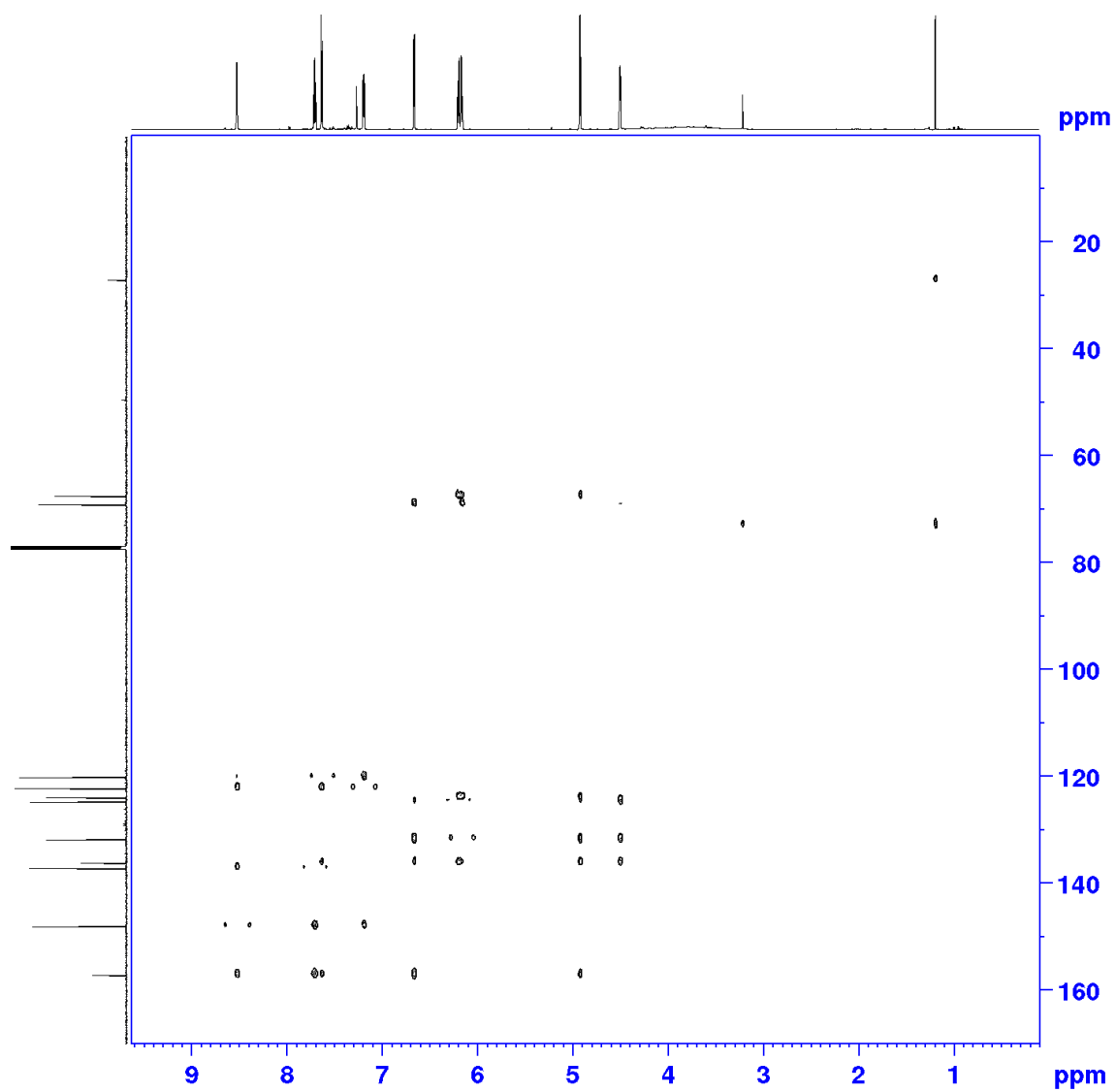Fig. S21: HSQC-NMR of (+)-(1*S*,2*R*)-3-(pyridin-2-yl)cyclohexa-3,5-diene-1,2-diol **3a** in CDCl₃.

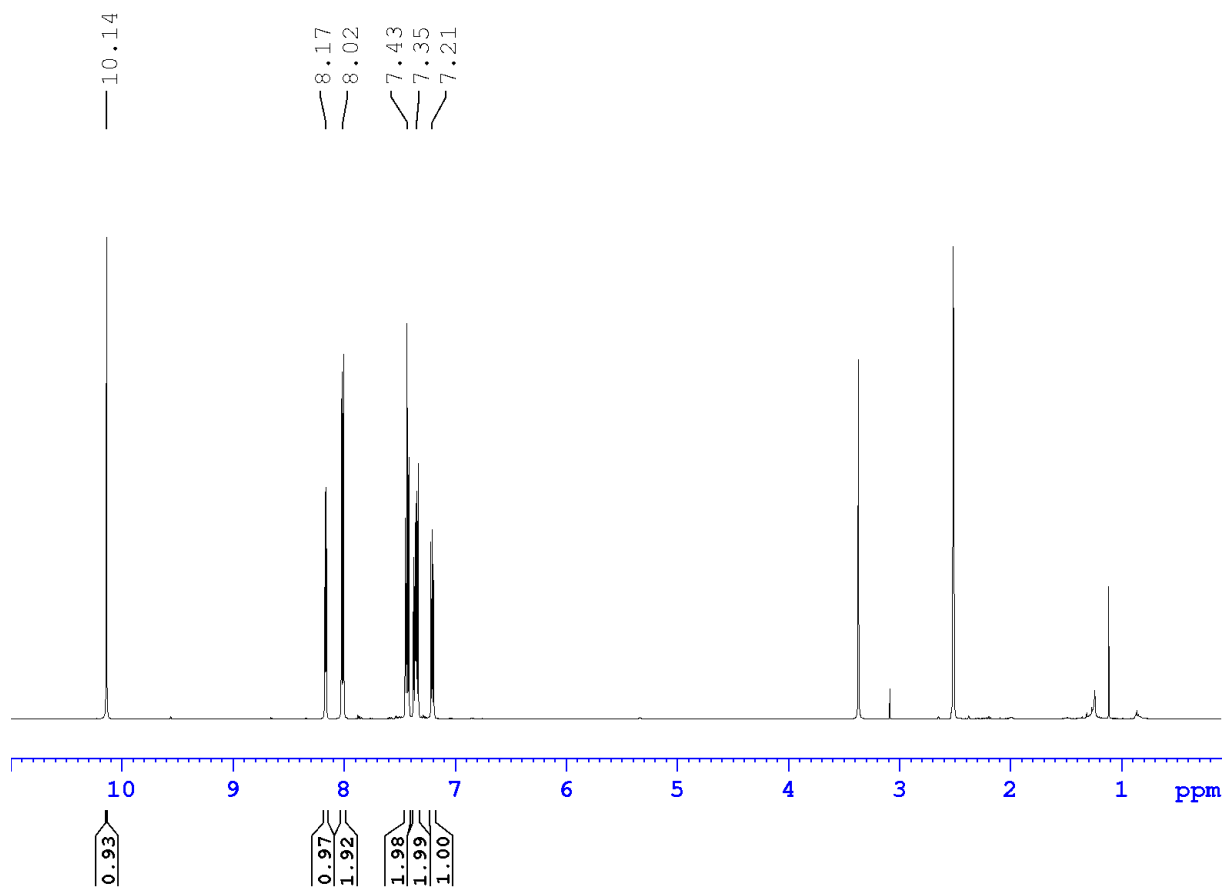Fig. S22: HMBC-NMR of (+)-(1*S*,2*R*)-3-(pyridin-2-yl)cyclohexa-3,5-diene-1,2-diol **3a** in CDCl$_3$.

Fig. S23: ^{1}H-NMR of 2-phenylpyridin-3-ol **3b** in DMSO-d6. H$_2$O and DMSO can be seen at 3.36 and 2.50 ppm, respectively.
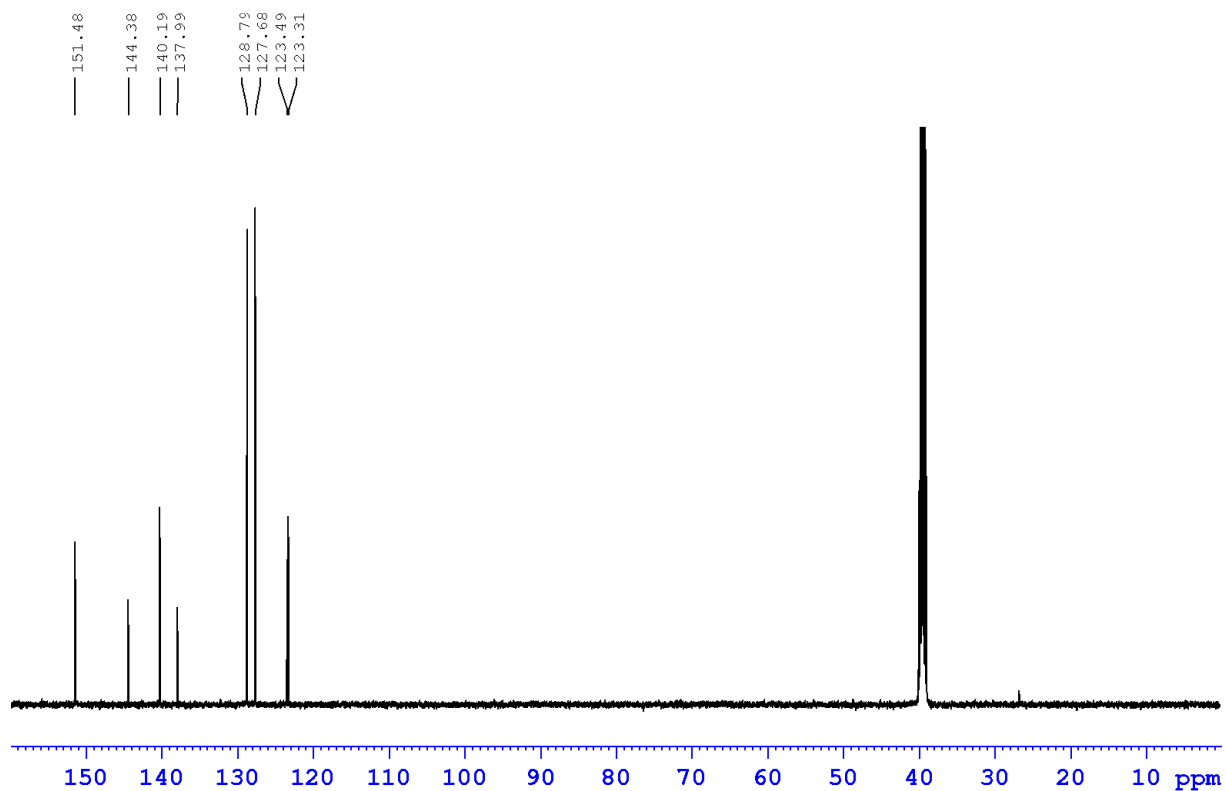

Fig. S24: ^{13}C-NMR of 2-phenylpyridin-3-ol **3b** in DMSO-d6. DMSO can be seen at 39.50 ppm.

Supplementary section 3: Semi-rational design of TDO variants

The 3DM information system of Bio-Prodict (Nijmegen, NL) was used for alignment of 671 ROs (stand: October 2018) of the three subfamilies; toluene/benzene, naphthalene, and cumene dioxygenase-like. A total of 27 different positions, consisting out of amino acids; 1) in the active site[1,2] 2) along the putative substrate entrance channel 3) previously described hot-spot positions in TDO or homologues, were chosen for site-directed mutagenesis (Table S1).
1) For the generation of active site variants, all positions with a conservation below 90% were chosen for mutagenesis. The 14 active site positions meeting this criterion, were mutated into amino acids, which exhibited, according to the 3DM sequence alignment, an occurrence of over 2% (Table S2).
2) The putative substrate channel was modeled *via* the CAVER PyMOL plugin according to supplementary section 6. To enlarge the substrate channel for the bulky bicyclic substrates **1**-**3**, six positions within a radius of 5 Å of the putative substrate channel were chosen to be exchanged to the small amino acid alanine (Table S3).
3) Seven positions, either previously described in literature or highlighted by the 3DM Bio-Prodict database as key positions (F114, T115, A212, A234, G232, V340 and T365), were chosen for mutagenesis; Positions F114 and T115 (corresponding to F99 and V100 in NDO)[3] are both located at the putative ferredoxin binding site and were emphasized by the 3DM Bio-Prodict database. A multiple TDO variant, containing mutations at the positions A212 and A234 exhibited an alteration in chemoselectivity and product formation for the substrate indene.[4] TDO position G323 corresponds to T335 in BPDO from *Burkholderia xenovorans* LB400, a hot-spot position influencing the catalytic activity.[5] According to the 3DM tool, variants at the TDO position V340 could exhibit an influence on active site residues. TDO position T365 was described previously to have a dramatic influence on the chemo- and enantioselectivity for the substrates propylbenzene, styrene and indene.[6]

Screening of the single point variants (Table S1) highlighted the positions M220 and A223, which showed the most tremendous increase in product formation, for all three substrates. Thus, they were chosen as template for the generation of a second mutant library consisting out of double variants (Table S4). To foster the product formation of **2a** and **3b** variants TDO$_{F114H\_A223T}$ and TDO$_{F216A\_F366V}$ were generated, respectively.

Table S1: Summary of all generated TDO single point variants (mutant library 1) with the corresponding 3DM numbers.

Position in TDO	3 DM number	Amino acid exchanges	Reason
F114	84	L, Y, H	Putative ferredoxin binding site[3]; hot-spot position according to 3DM
T115	85	A, V	Putative ferredoxin binding site[3]; hot-spot position according to 3DM
A212	177	T	Previously described[4]
Q215	180	N, G	Active site
F216	181	A, G, I, L, N, Q, S, T, V, Y	Active site
M220	185	A, C, D, F, G, I, L, N, P, Q, S, T, V, W	Active site
A223	188	C, E, F, G, I, K, L, M, N, P, Q, S, T, V, W	Active site

17

G224	-	A	Substrate channel
T225	-	A	Substrate channel
L229	193	A	Substrate channel
A234	198	F, S	Previously described[4]
L245	-	A	Substrate channel
P247	-	A	Substrate channel
P248	-	A	Substrate channel
G264	227	A, L, M, N, S	Active site
Y266	229	F, K, L, M, N, S	Active site
L272	-	A, G	Active site
I276	-	A, G	Active site
V309	232	A, F, G, H, N, S	Active site
H311	234	A, I, M, N, S, V, Y	Active site
L321	244	A, F, N, S	Active site
G323	246	A, C, Q	Previously described[5]
I324	247	A, F, G, M, N, S, T	Active site
V340	263	A, F	Hot-spot position according to 3DM
T365	288	A, F, L, N	Previously described[6]
F366	289	A, I, L, N, Q, S, T, V, Y	Active site
F372	293	A, G, L, N, Q, S, T, W, Y	Active site

Table S2: All TDO active site positions with the corresponding variants generated in this study.[1,2]

Position in TDO	Amino acid exchange
Q215	N, G
F216	A, G, I, L, N, Q, S, T, V, Y
[a]D219	-
M220	A, C, D, F, G, I, L, N, P, Q, S, T, V, W
[a]H222	-
A223	C, E, F, G, I, K, L, M, N, P, Q, S, T, V, W
[a]H228	-
G264	A, L, M, N, S
Y266	F, K, L, M, N, S
L272	A, G
I276	A, G
V309	A, F, G, H, N, S
H311	A, I, M, N, S, V, Y
L321	A, F, N, S
I324	A, F, G, M, N, S, T
F366	A, I, L, N, Q, S, T, V, Y
F372	A, G, L, N, Q, S, T, W, Y
[a]D376	-

a: positions H222, H228 and D376 were not considered for mutagenesis, since all three are coordinating the catalytic iron and are exhibiting conservations above 90%. Active site position D219 was also not considered due to its high conservation (99%).

Table S3: TDO positions located along the putative substrate entrance channel, outside the active pocket. Amino acids within radius of 5 Å of the simulated channel (CAVER PyMOL plugin) were replaced by alanine.

| Position in TDO	Amino acid exchange
G224	A
T225	A
L229	A
[a]I232	-
L245	A
P247	A
P248	A

a: Variant L232A could not be generated

Table S4: All 55 generated TDO double variants (mutant library 2).

TDO double variant	TDO double variant
F114H_A223T	M220C_A223P
F216A_F366V	M220C_A223Q
M220A_A223C	M220C_A223S
M220A_A223E	M220C_A223T
M220A_A223G	M220C_A223V
M220A_A223I	M220C_A223W
M220A_A223K	M220C_V309G
M220A_A223L	M220C_L321A
M220A_A223M	M220S_A223C
M220A_A223P	M220S_A223G
M220A_A223Q	M220S_A223E
M220A_A223S	M220S_A223I
M220A_A223T	M220S_A223K
M220A_A223V	M220S_A223L
M220A_A223W	M220S_A223M
M220A_V309G	M220S_A223P
M220A_L321A	M220S_A223Q
M220A_F366L	M220S_A223S
M220A_F366Q	M220S_A223T
M220A_ F366T	M220S_A223V
M220A_F366V	M220S_A223W
M220C_A223C	M220S_L272G
M220C_A223E	M220S_Y266N
M220C_A223G	A223T_V309G
M220C_A223I	A223T_L321A
M220C_A223K	A223V_L321A
M220C_A223L	A223V_V309G
M220C_A223M	

Supplementary section 4: Screening data

4.1 Screening results for naphthalene **1**
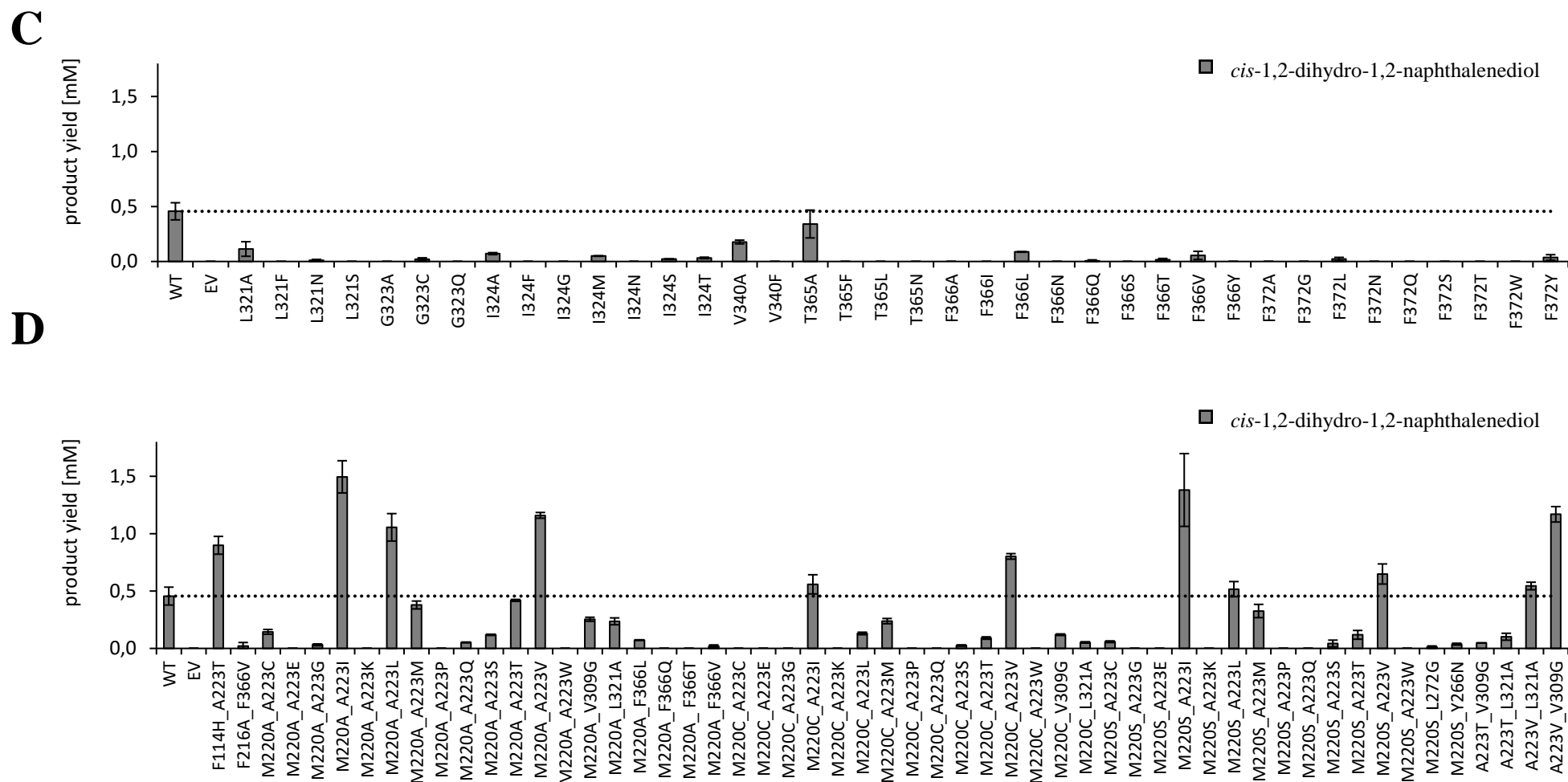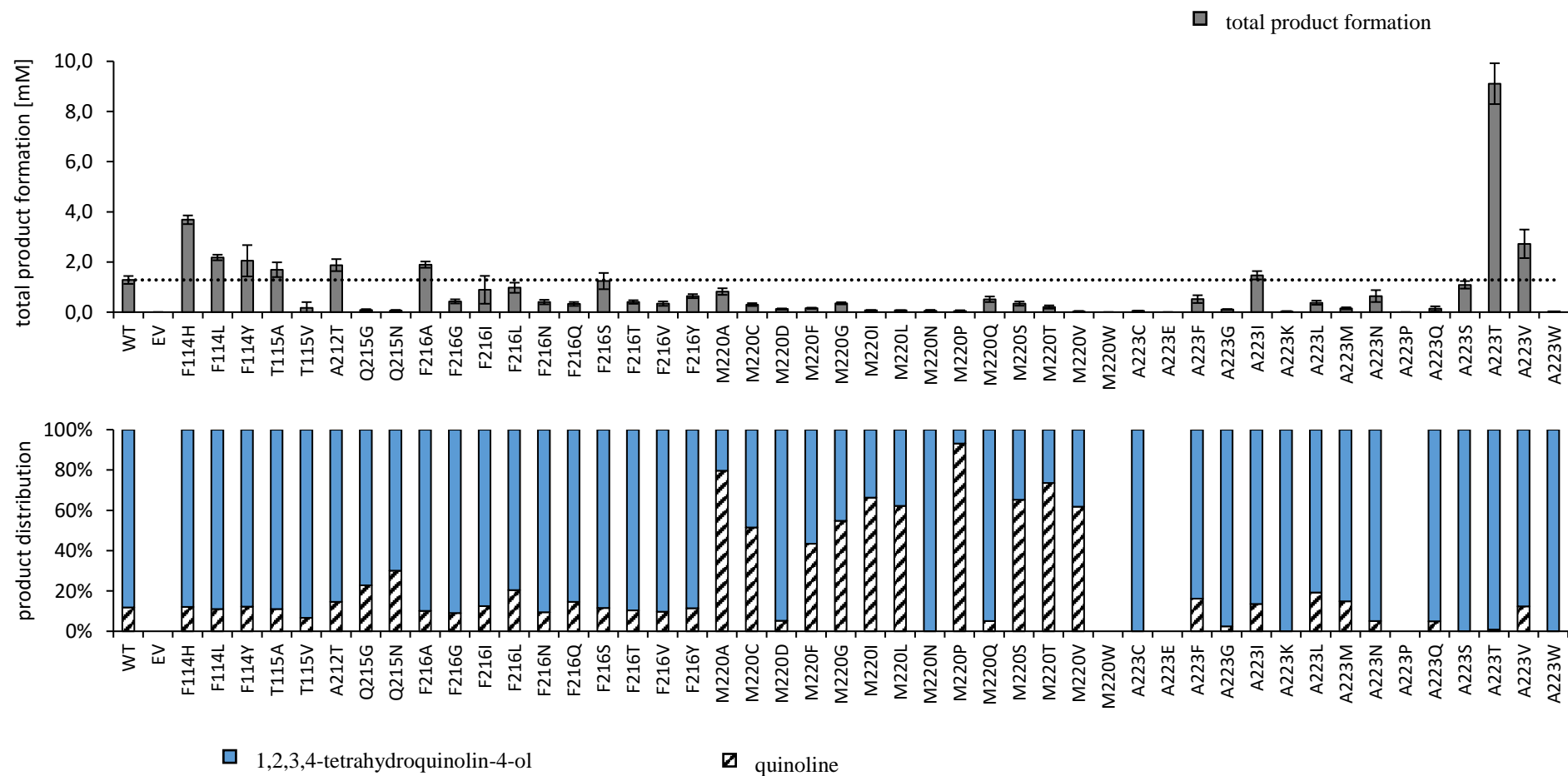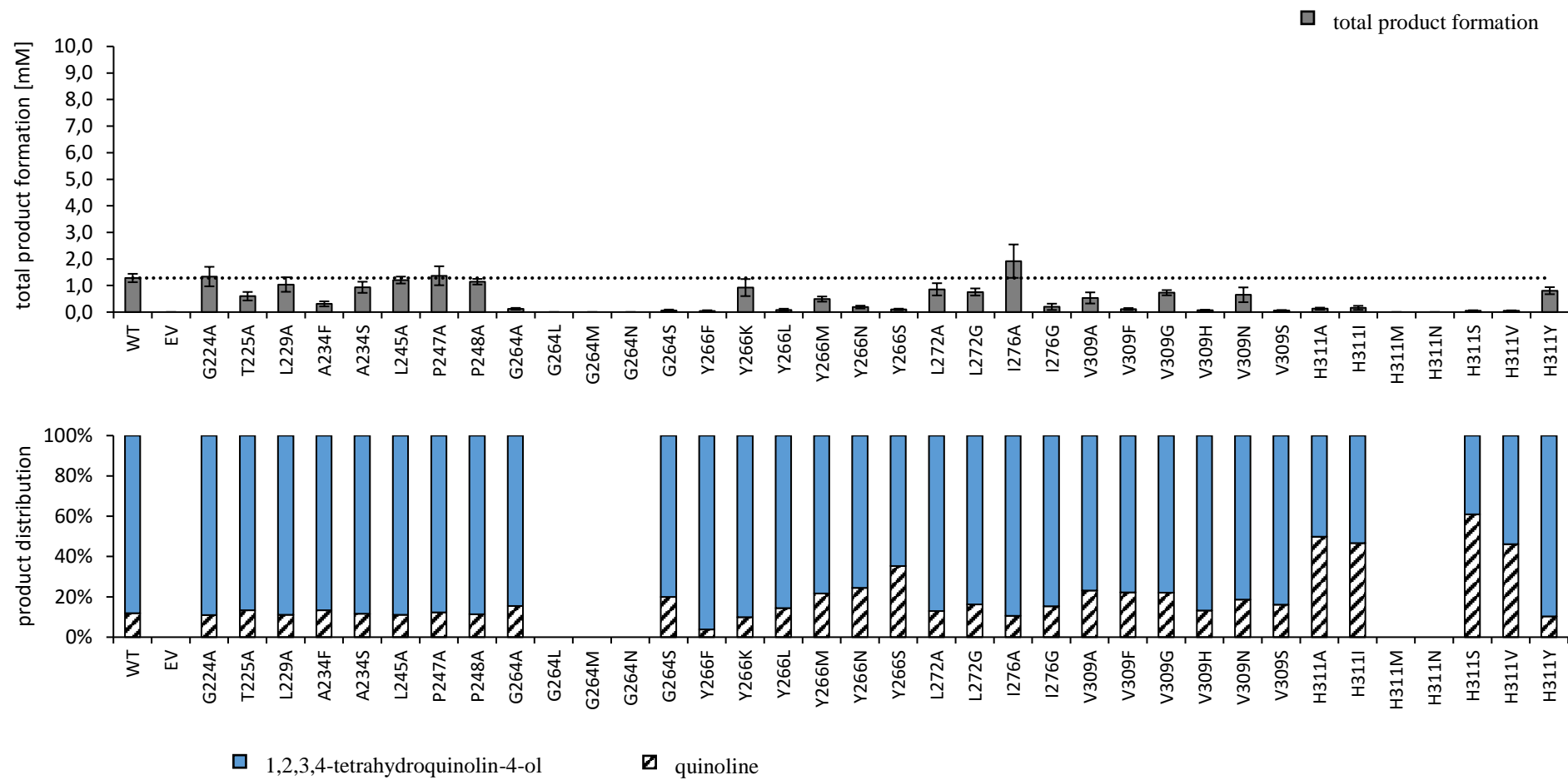
A

B

C

D
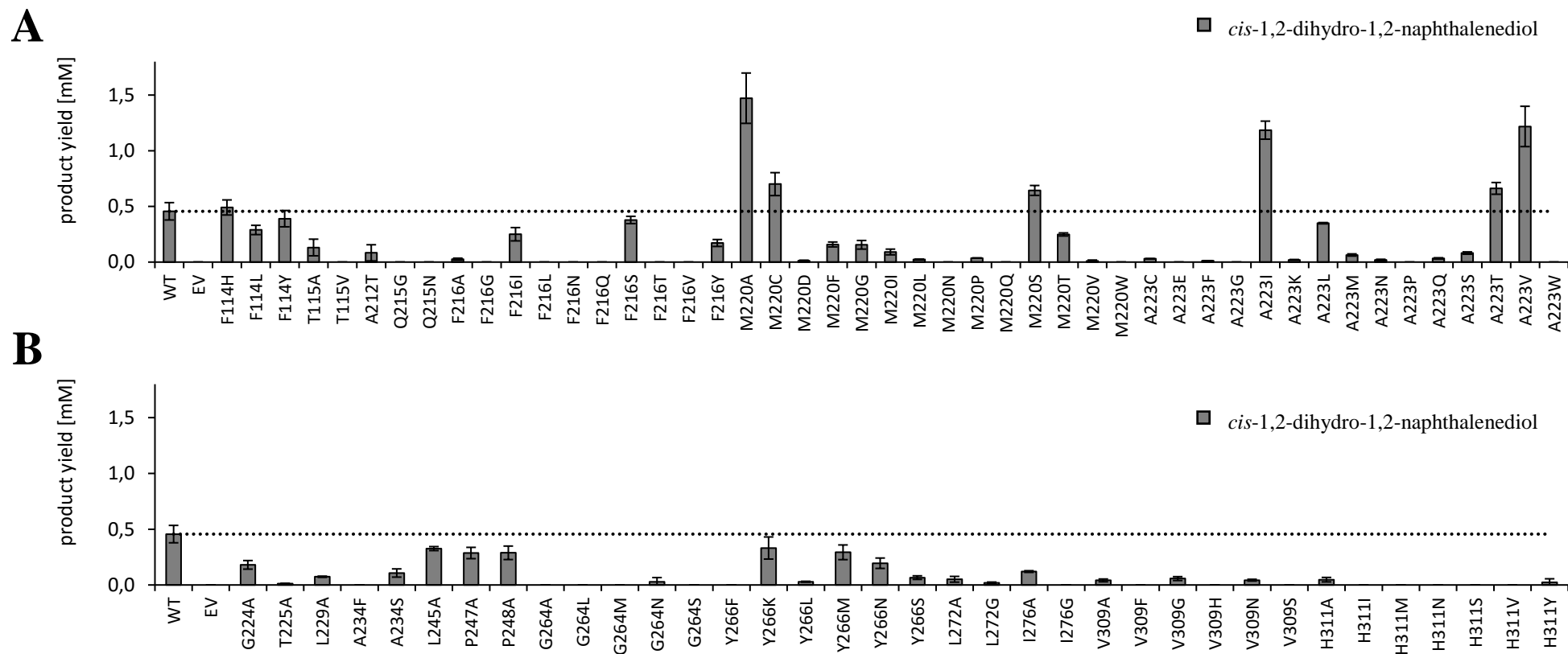
Fig. S25: Screening of mutant library 1 and 2 with the substrate naphthalene **1**. *cis*-1,2-Dihydro-1,2-naphthalenediol **1a** formation for each tested variant is shown in grey. TDO wild type product formation is displayed as a dotted black line. Biotransformations were performed according to experimental section 1.3. Each bar is showing the mean value of three biological triplicates and the corresponding standard deviations. **A**: Wild type, empty vector, and variants F114H to A223W; **B**: Wild type, empty vector, and variants G224A to H311Y; **C**: Wild type, empty vector, and variants L321A to F372Y; **D**: Wild type, empty vector, and double variants F114H_A224T to A223V_V309G.
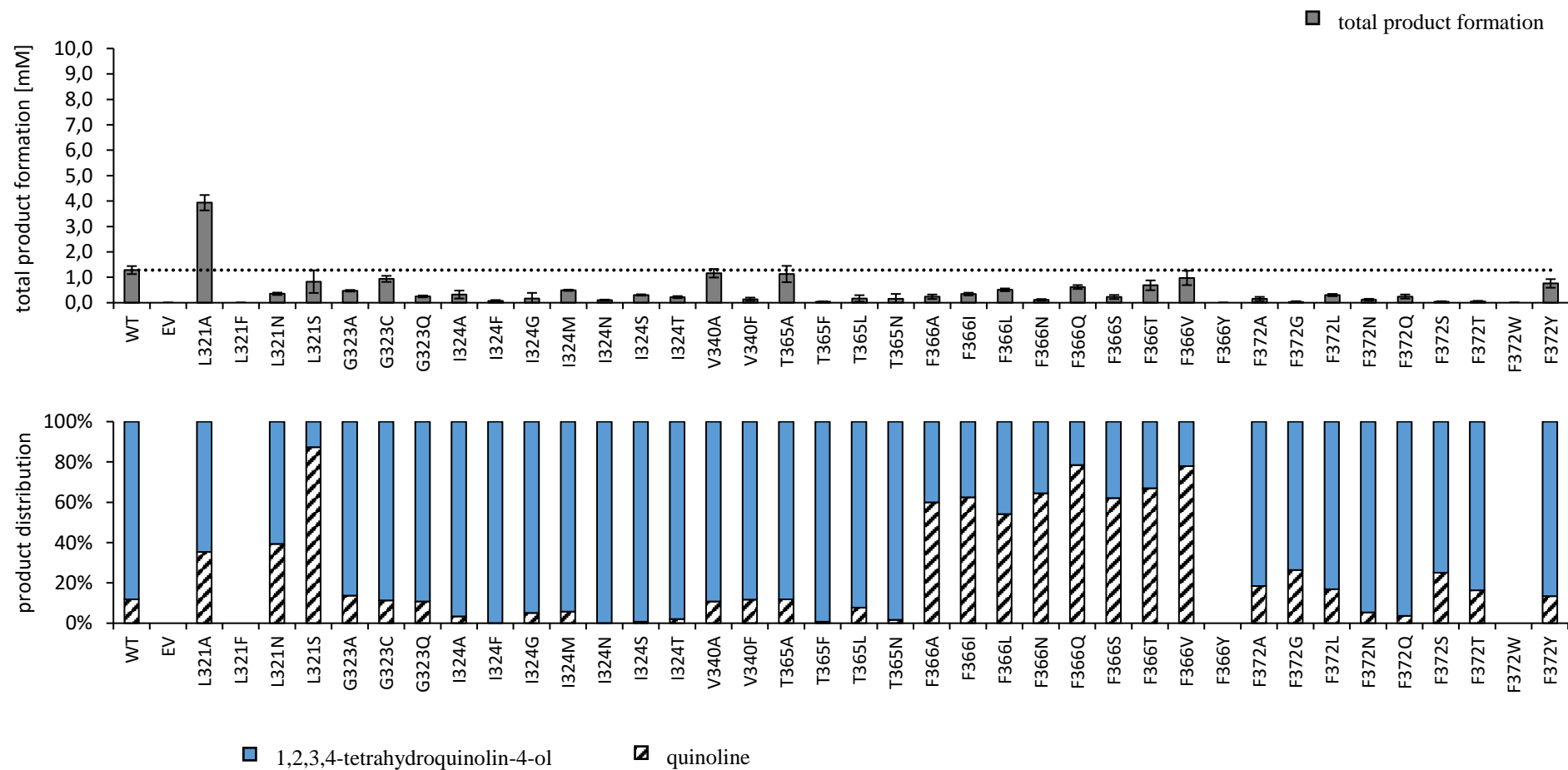
4.2 Screening results for 1,2,3,4-tetrahydroquinoline **2**
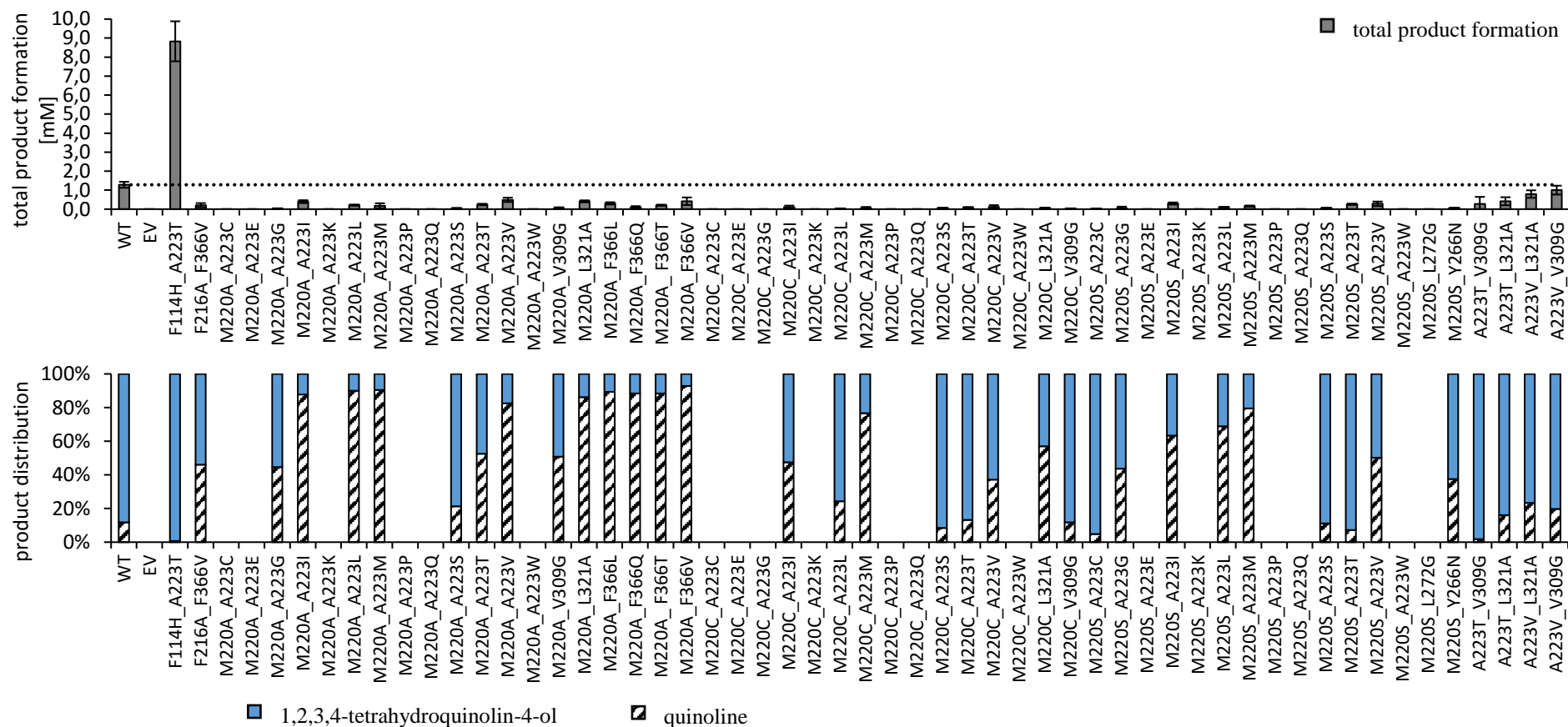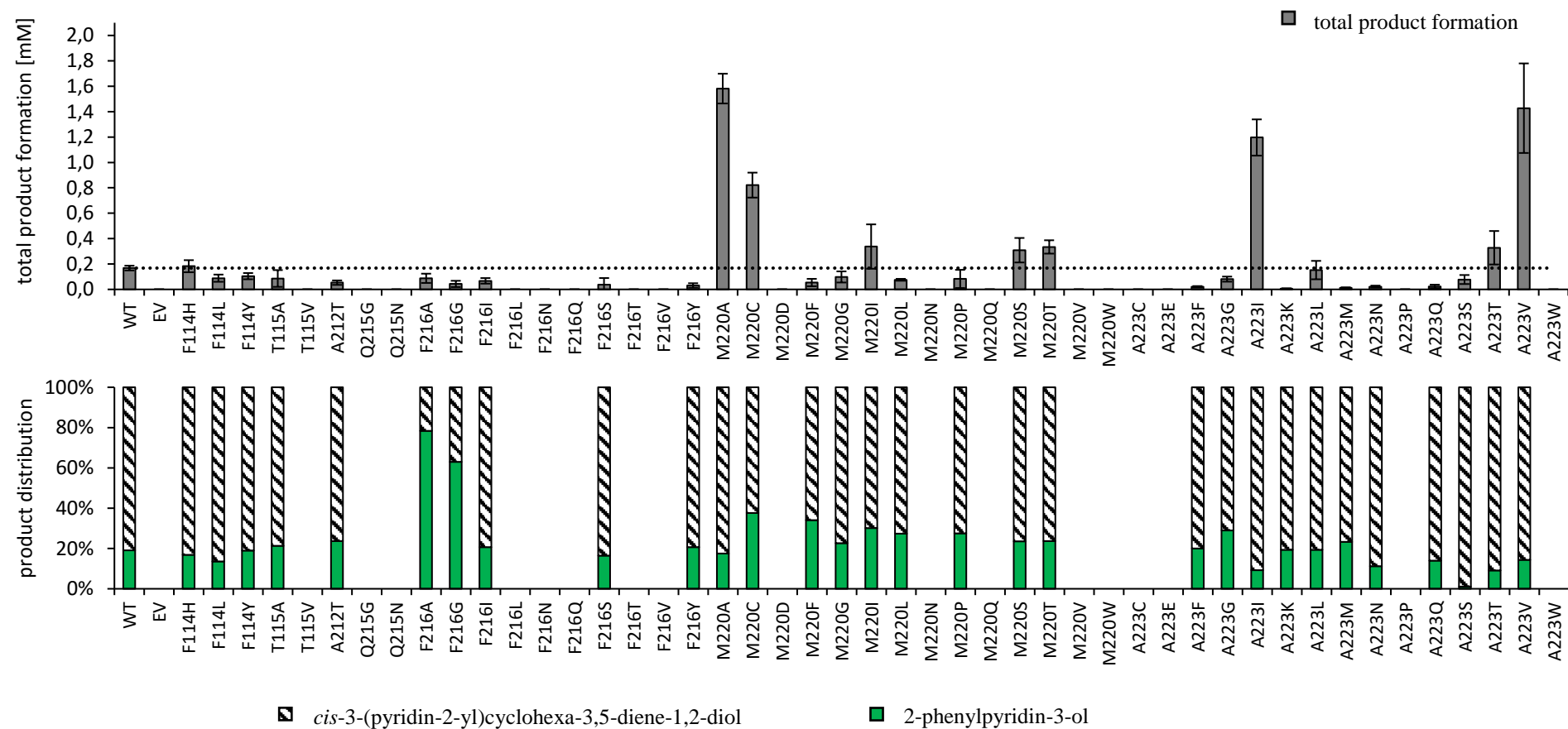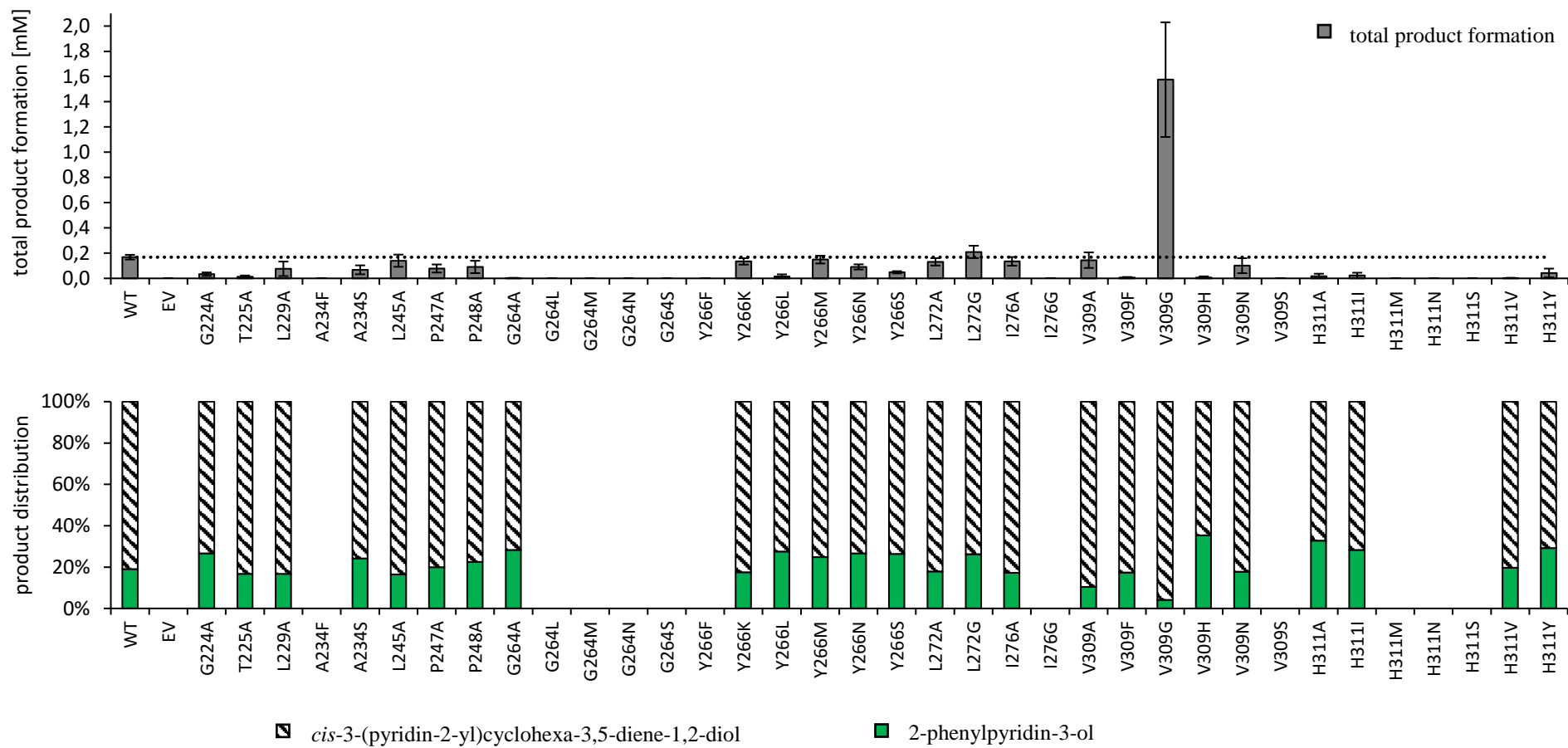
A

B

C

24

D

Fig. S26: Screening of mutant library 1 and 2 with the substrate 1,2,3,4-tetrahydroquinoline **2**. Total product formation for each tested variant is shown in grey (top chart). TDO wild type product formation is displayed as a dotted black line. Each bar is showing the mean value of three biological triplicates and the corresponding standard deviations. The product distribution of 1,2,3,4-tetrahydroquinoline-4-ol **2a** (blue) to quinoline **2b** (black striped) is depicted below the total product formation for each tested variant (bottom chart). Biotransformations were performed according to experimental section 1.3. **A**: Wild type, empty vector, and variants F114H to A223W; **B**: Wild type, empty vector, and variants G224A to H311Y; **C**: Wild type, empty vector, and variants L321A to F372Y; **D**: Wild type, empty vector, and double variants F114H_A224T to A223V_V309G.
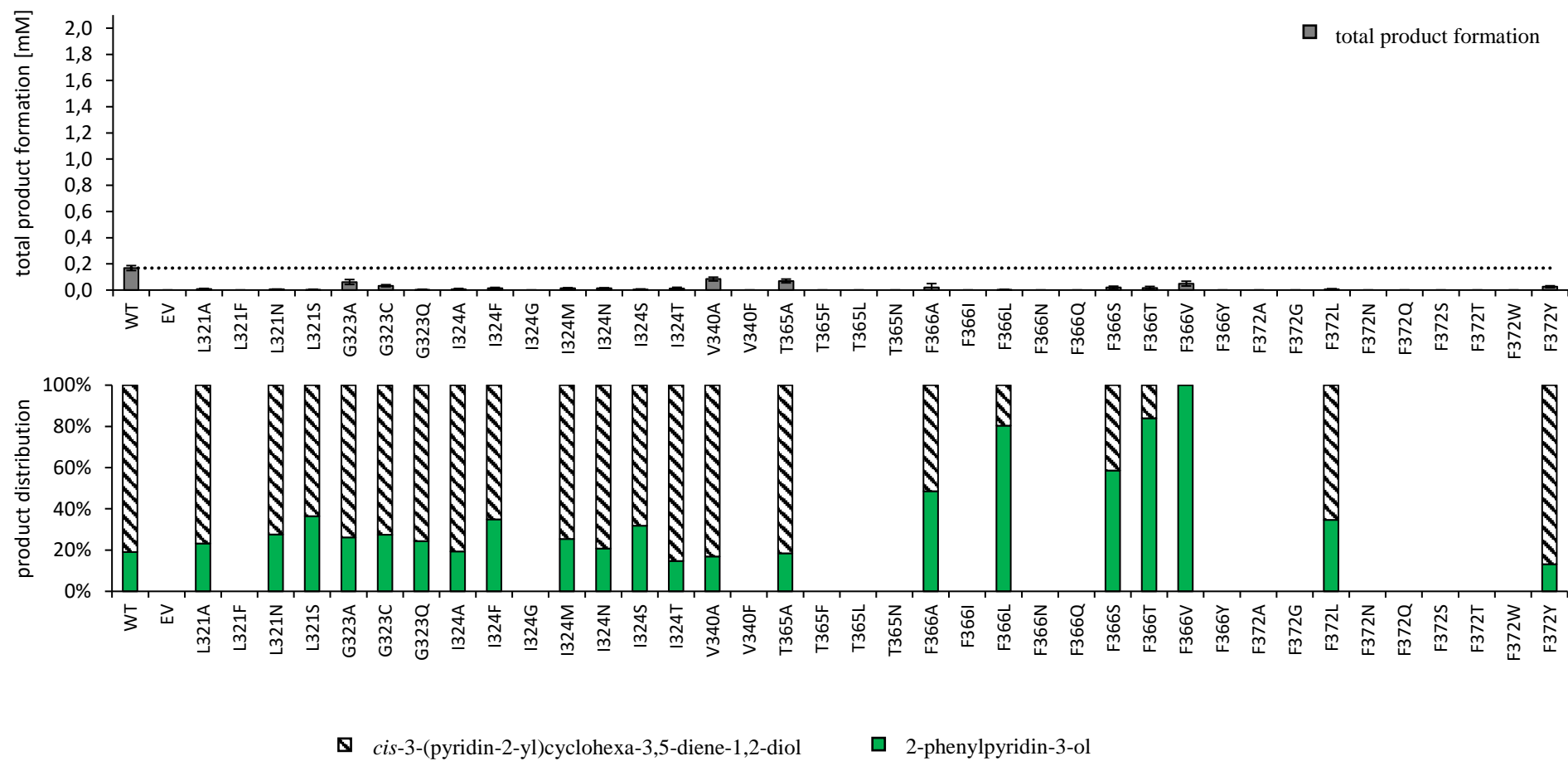
4.3 Screening results for 2-phenylpyridine 3

A

cis-3-(pyridin-2-yl)cyclohexa-3,5-diene-1,2-diol   2-phenylpyridin-3-ol

B

cis-3-(pyridin-2-yl)cyclohexa-3,5-diene-1,2-diol   2-phenylpyridin-3-ol

C

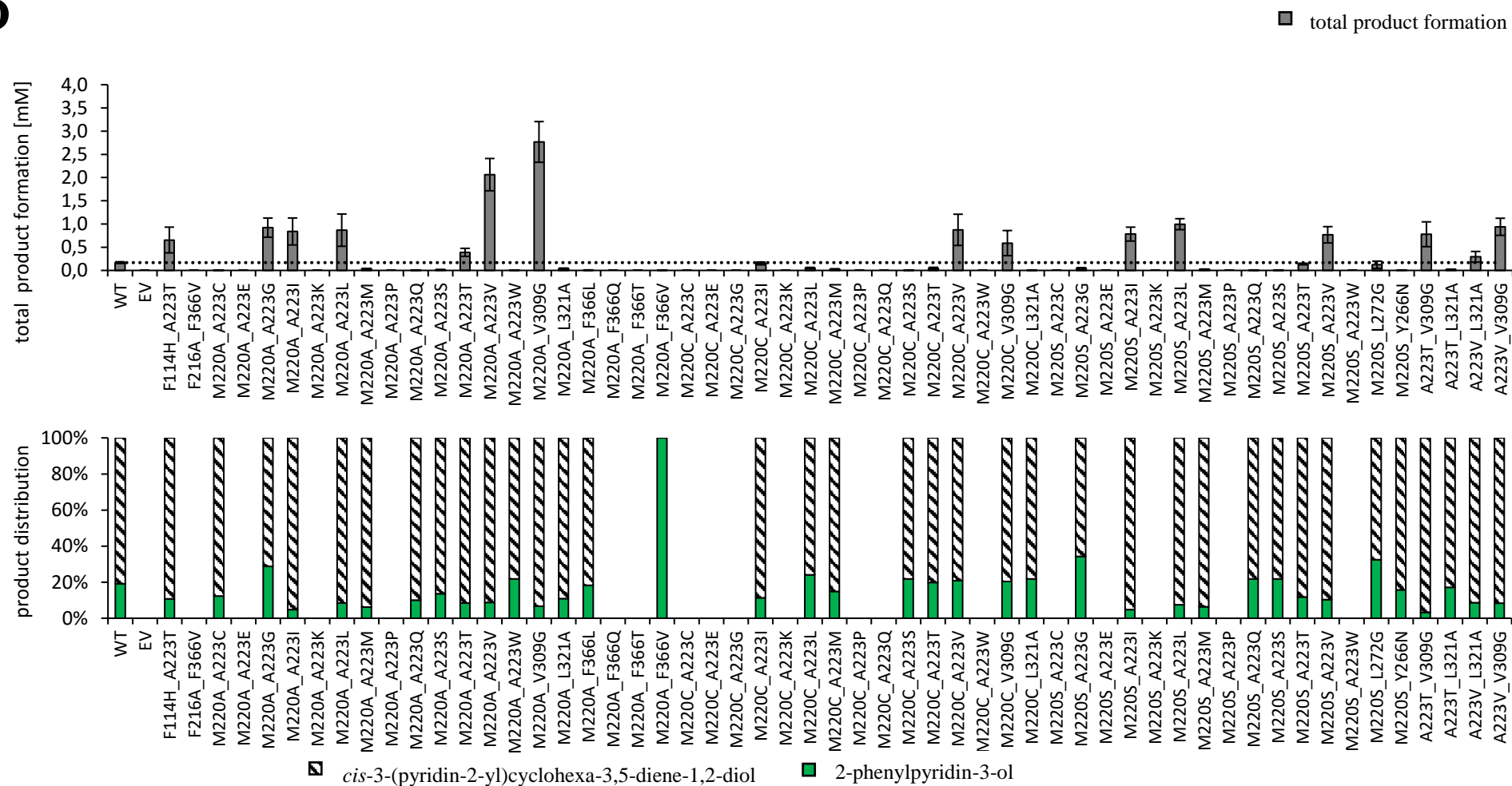cis-3-(pyridin-2-yl)cyclohexa-3,5-diene-1,2-diol   2-phenylpyridin-3-ol

D

Fig. S27: Screening of mutant library 1 and 2 with the substrate 2-phenylpyridine **3**. Total product formation for each tested variant is shown in grey (top chart). TDO wild type product formation is displayed as a dotted black line. Each bar is showing the mean value of three biological triplicates and the corresponding standard deviations. The product distribution of *cis*-3-(pyridin-2-yl)cyclohexa-3,5-diene-1,2-diol **3a** (black striped) to 2-phenylpyridin-3-ol **3b** (green) is depicted below the total product formation for each tested variant (bottom chart). Biotransformations were performed according to experimental section 1.3. **A**: Wild type, empty vector, and variants F114H to A223W; **B**: Wild type, empty vector, and variants G224A to H311Y; **C**: Wild type, empty vector, and variants L321A to F372Y; **D**: Wild type, empty vector, and double variants F114H_A224T to A223V_V309G.

Supplementary section 5: pH effect influencing the enantiomeric excess for 2a

Zheng and colleagues showed that the pH has a strong influence in the enantiomeric excess of (+)-(*R*)-1,2,3,4-tetrahydroquinoline-4-ol **2a**, when biosynthesized with the wild type strain *Pseudomonas plecoglossicidas*.[7] Lower pH-values led to an decrease in enantiomeric excess, due to racemization of (+)-(*R*)-1,2,3,4-tetrahydroquinoline-4-ol **2a** at acidic conditions. To increase the enantiomeric excess of **2a**, we investigated the impact of the chosen pH and reaction buffer (Fig. S28).
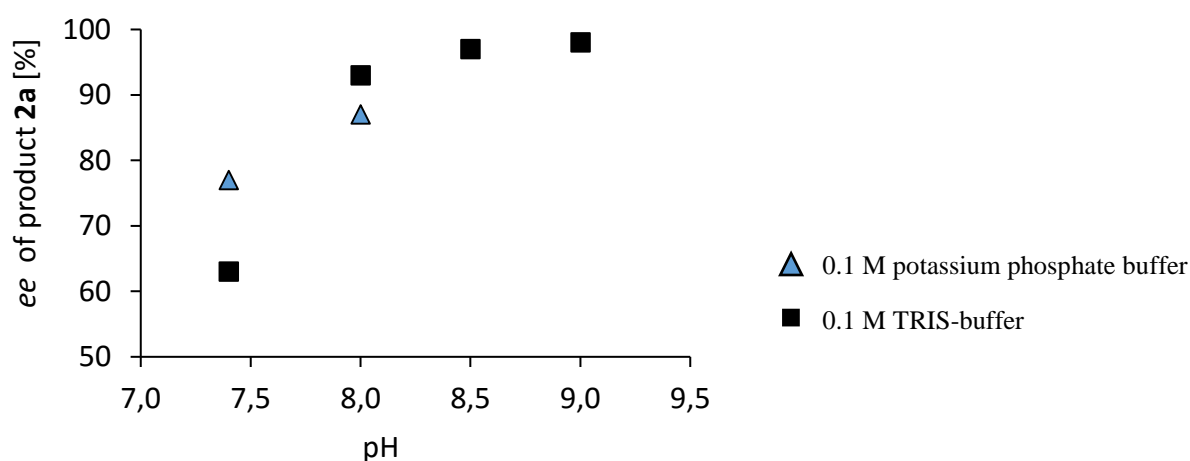

Fig. S28: Enantiomeric excess of (+)-(*R*)-1,2,3,4-tetrahydroquinoline-4-ol **2a,** biosynthesized with TDO wild type, employing 1,2,3,4-tetrahydroquinoline **2** as substrate at different pH-values. Two different reaction buffers were tested; 0.1 M potassium phosphate pH 7.4 and 8.0 (blue triangle) and 0.1 TRIS-buffer pH 7.4, 8.0, 8.5, and 9.0 (black square). Reactions were performed in 96 deep well plates as described in experimental section 1.4.

As expected, the enantiomeric excess for **2a** increased at higher pH values. Biotransformations with 0.1 M TRIS buffer at pH 9.0 showed an *ee* of >98% but a decreased product formation was observed (data not shown). Biotransformations performed with 0.1 M TRIS buffer at pH 8.5, resulted in an advantageous *ee* of 97%, compared to the starting conditions (0.1 M potassium phosphate buffer pH 7.4), while maintaining the same level of product formation (data not shown). We carried out biotransformations with selected outperforming variants favoring the production of (+)-(*R*)-1,2,3,4-tetrahydroquinoline-4-ol **2a,** according to experimental section 1.4. Product formation and the enantiomeric excess were measured for each case (Table S8).

Table S8: Validation of variants with increased product formation and higher selectivity for (+)-(*R*)-1,2,3,4-tetrahydroquinoline-4-ol **2a**. Biotransformations were performed at pH 8.5 with 0.1 M TRIS-buffer in biological triplicates.

Variant	Product formation of 2a [mM]	Product formation of 2b [mM]	Enantiomeric excess of 2a [%]
WT	3.82±0.44	0.32±0.02	97
F114H	5.86±0.03	0.44±0.01	96
A223T	7.87±0.08	0.01±0.01	93
F114H_A223T	8.57±0.18	0.01±0.01	95

Double variant TDO$_{F114H\_A223T}$ outperformed over the rest of the tested variants, enabling the selective production of 8.58 mM (86%) (+)-(*R*)-1,2,3,4-tetrahydroquinoline-4-ol **2a** with an *ee* value of 95% and was chosen for semi-preparative biotransformation (experimental section 1.5).

Supplementary section 6: Computational visualization of TDO active site and substrate channel

The crystal structure of the TDO α-subunit from *P. putida* F1 (PDB ID; 3EN1)[2] was employed for the visualization of the active site and for modeling the substrate channel in PyMOL (Schrödinger Inc., New York, US; PyMOL version 2.4). All 18 active site amino acids (Table S2) were employed for the illustration of the active site cavity (cavity detection radius; 3 solvent radii, cavity detection cutoff; 1 solvent radius). Locations of six active site TDO positions with major influence on product formation and chemo-, regio- and enantioselectivity for bicyclic compounds (Table 4) are shown in Fig. S29.
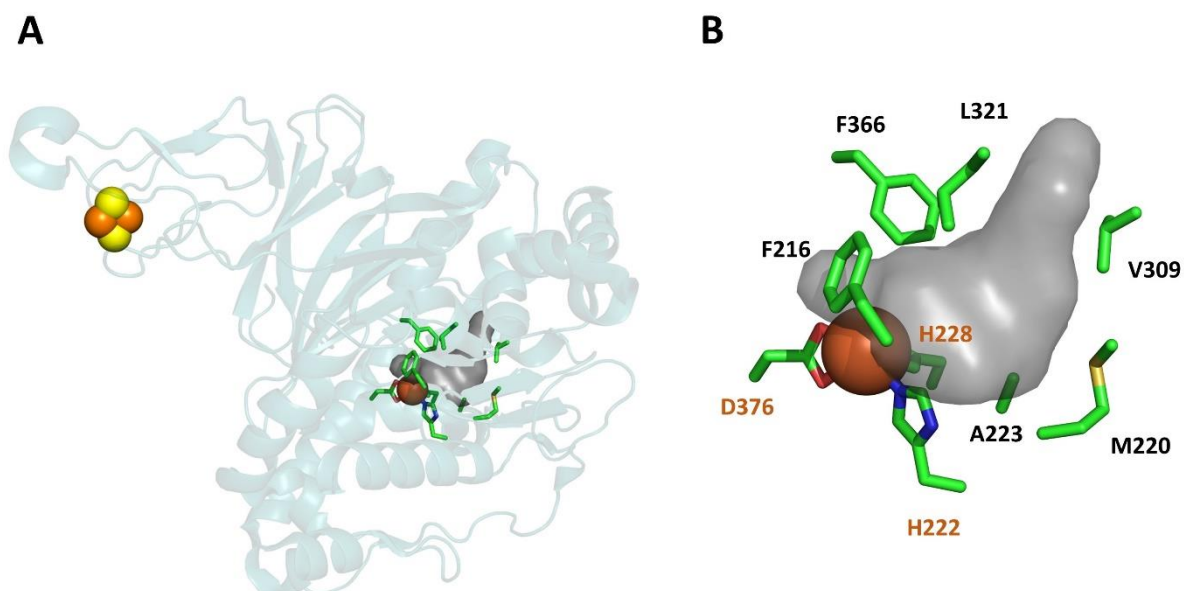

Fig. S29: **A** TDO wild type α-subunit active pocket (grey), surrounded by six active site hot-spot positions (green). The Rieske center (2Fe-2S) can be seen as spheres on the left; sulfur (in yellow) and the catalytic active iron located at the active pocket (in orange). **B** Close up of the active pocket. Residues of the six hot-spot positions are depicted according the following color code; light green; C, red; O, blue; N, yellow; S. To facilitate orientation at the active pocket, active site residues H222, H228 and D376, coordinating the catalytic iron, are included and highlighted in orange. Hot-spot position F114 cannot be visualized, since is not located at the active site.

The substrate channel was modeled using the CAVER plugin[8] for PyMOL (CAVER version 3.0.1). As starting point for the calculations, the catalytically active iron was resolved. The following standard parameters were chosen; minimum probe radius 0.9, shell depth 4, shell radius 3, clustering threshold 3.5, maximum distance 3 Å, and desired radius 5 Å (Fig. S30).
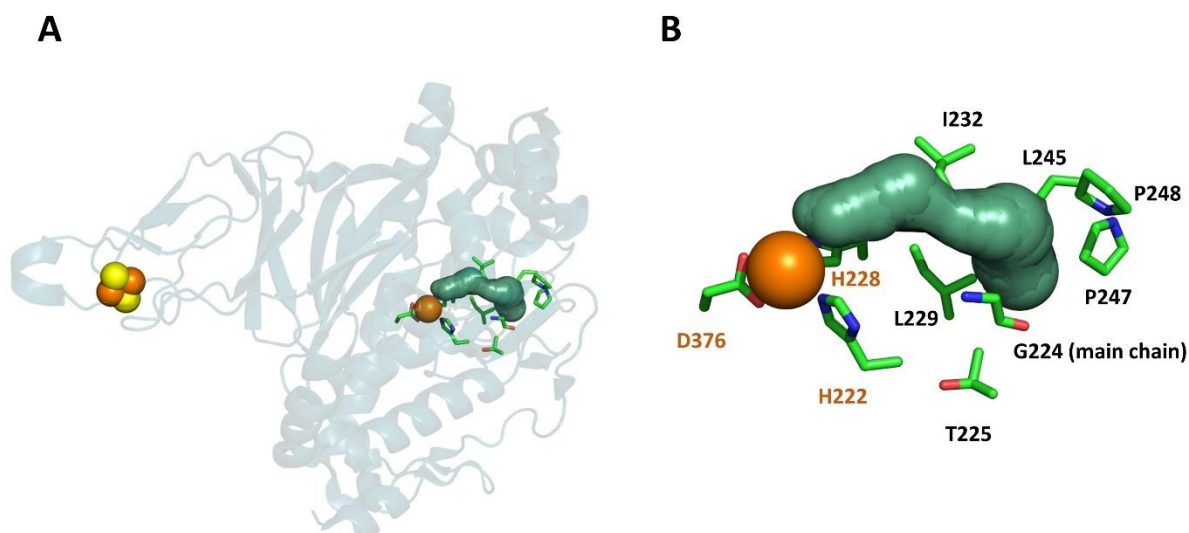
Fig. S30: **A** Putative substrate channel (dark green) envisaged for TDO wild type α-subunit, computed with CAVER. The Rieske center (2Fe-2S) can be seen as spheres on the left; sulfur (in yellow) and the catalytic active iron located at the active site (in orange). The later, was employed as starting point to perform CAVER calculations. **B** Close up of the substrate channel. Seven side chain residues within 5 Å proximity of the substrate channel, but not located in the active site, are depicted (light green; C, red; O, blue; N). To facilitate orientation at the active pocket, active site residues H222, H228 and D376, coordinating the catalytic iron, are included and highlighted in orange. In addition, for residue G224 the main chain is provided.

Structural models were generated in order to compare the size of the TDO wild-type active site with TDO$_{M220A\_V309G}$, which was the outperforming variant enabling the highest conversion of the bulky substrate 2-phenylpyridine **3**. For model generation, the amino acid sequence of TDO$_{M220A\_V309G}$ was computed, using the protein structure homology-modelling platform SWISS-MODEL[9] (Biozentrum, University of Basel, CH), employing TDO wild-type sequence as template (PDB ID; 3EN1). A comparative view of the generated homology models highlighted an expansion in the active site cavity of variant TDO$_{M220A\_V309G}$, in comparison to TDO wild-type (Fig. S31).
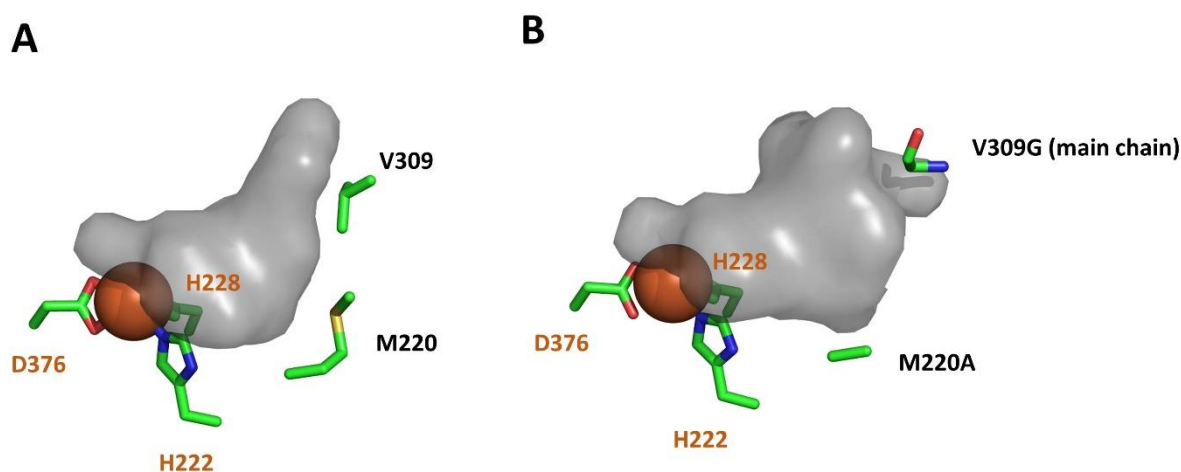

Fig. S31: Generated active site models for TDO wild-type **A** and the double variant TDO$_{M220A\_V309G}$ **B**. In both cases, the active site is shown in grey, and active site positions M220 and V309 are included in the illustration. Amino acid residues are depicted according to the following color code; light green; C, red; O, blue; N, yellow; S. To facilitate orientation at the active pocket, active site residues H222, H228 and D376, coordinating the catalytic iron, are included and highlighted in orange. In addition, for residue V309G the main chain is provided.

32

References

[1]   M. A. Vila, D. Umpiérrez, G. Seoane, S. Rodríguez, I. Carrera, N. Veiga, *J. Mol. Catal. B Enzym.* **2016**, *133*, S410–S419.

[2]   R. Friemann, K. Lee, E. N. Brown, D. T. Gibson, H. Eklund, S. Ramaswamy, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2009**, *65*, 24–33.

[3]   B. Kauppi, K. Lee, E. Carredano, R. E. Parales, D. T. Gibson, H. Eklund, S. Ramaswamy, *Structure* **1998**, *6*, 571–586.

[4]   N. Zhang, B. G. Stewart, J. C. Moore, R. L. Greasham, D. K. Robinson, B. C. Buckland, C. Lee, *Metab. Eng.* **2000**, *2*, 339–348.

[5]   M. Mohammadi, J. F. Viger, P. Kumar, D. Barriault, J. T. Bolin, M. Sylvestre, *J. Biol. Chem.* **2011**, *286*, 27612–27621.

[6]   M. A. Vila, D. Umpiérrez, N. Veiga, G. Seoane, I. Carrera, S. Rodríguez Giordano, *Adv. Synth. Catal.* **2017**, *359*, DOI 10.1002/adsc.201700444.

[7]   D. Zheng, M. Yang, J. Zhuo, K. Li, H. Zhang, J. Yang, B. Cui, Y. Chen, *J. Mol. Catal. B Enzym.* **2014**, *110*, 87–91.

[8]   E. Chovancova, A. Pavelka, P. Benes, O. Strnad, J. Brezovsky, B. Kozlikova, A. Gora, V. Sustr, M. Klvana, P. Medek, L. Biedermannova, J. Sochor, J. Damborsky, *PLoS Comput. Biol.* **2012**, *8*, 23–30.

[9]   A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F. T. Heer, T. A. P. De Beer, C. Rempfer, L. Bordoli, R. Lepore, T. Schwede, *Nucleic Acids Res.* **2018**, *46*, W296–W303.